

Utah State University

DigitalCommons@USU

All Graduate Theses and Dissertations

Graduate Studies

5-1997

An Investigation of the Temporal Stability of Self-Reported Internalizing Symptoms in Elementary-Age Children

Kurt David Michael
Utah State University

Follow this and additional works at: <https://digitalcommons.usu.edu/etd>



Part of the [Psychology Commons](#)

Recommended Citation

Michael, Kurt David, "An Investigation of the Temporal Stability of Self-Reported Internalizing Symptoms in Elementary-Age Children" (1997). *All Graduate Theses and Dissertations*. 6106.

<https://digitalcommons.usu.edu/etd/6106>

This Thesis is brought to you for free and open access by the Graduate Studies at DigitalCommons@USU. It has been accepted for inclusion in All Graduate Theses and Dissertations by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.



AN INVESTIGATION OF THE TEMPORAL STABILITY OF
SELF-REPORTED INTERNALIZING SYMPTOMS
IN ELEMENTARY-AGE CHILDREN

by

Kurt David Michael

A thesis submitted in partial fulfillment
of the requirements for the degree

of

MASTER OF SCIENCE

in

Psychology

Approved:

UTAH STATE UNIVERSITY
Logan, Utah

1997

Copyright © Kurt David Michael 1997

All Rights Reserved

ABSTRACT

An Investigation of the Temporal Stability of
Self-Reported Internalizing Symptoms
in Elementary-Age Children

by

Kurt David Michael, Master of Science

Utah State University, 1997

Major Professor: Dr. Kenneth W. Merrell
Department: Psychology

Over the past two decades, a great deal of research has been devoted to the understanding of internalizing disorders in children. Internalizing disorders encompass a wide variety of problems, including depression, anxiety, social withdrawal, and somatic complaints. It has been suggested that the existence of internalizing disorders in children has negative effects upon their self-esteem, academic achievement, physical health, and future adjustment. However, because internalizing disorders are, in great measure, subjective perceptions of internal distress, they are often not readily or reliably identified by external observers. As a result, several researchers have stressed the importance of eliciting the child's perspective through self-report assessment. While there are several excellent self-report measures of internalizing constructs, none of these instruments is designed to measure the comprehensive domain of internalizing disorders in children

below the age of 11 even though it has been established that children as young as 8 are able to give reliable self-reports. This apparent dearth of broad-based instruments for middle- to late-elementary school children creates problems for the assessment of internalizing problems because the various internalizing syndromes often coexist with one another, therefore limiting the utility of a single-syndrome instrument.

The newly developed Internalizing Symptoms Scale for Children (ISSC) is a 48-item self-report instrument designed to measure the broad range of internalizing problems in children. This investigation was conducted to establish whether the ISSC is a reliable measure of internalizing symptoms in 8- to 12-year-old children over 2-, 4-, and 12-week intervals. Overall, the findings provide strong support for the ISSC as a reliable measure of internalizing symptoms in elementary-age children over short- to medium-length time intervals.

(110 pages)

DEDICATION

I dedicate this thesis to my wife, Amy, and my newborn son, Kauner. I extend my most sincere acknowledgment to my wife, Amy, for her unwavering personal support, spiritual guidance, and unconditional love in life as well as during the pursuit of my professional aspirations. Her ability to withstand my efforts long into the night, my frustration, and my mood swings did not go unnoticed. Finally, her recognition of my personal triumphs is greatly appreciated.

ACKNOWLEDGMENTS

I want to express my deep appreciation to Dr. Kenneth Merrell for his support and guidance throughout this thesis project. From the project's inception to the final defense, Dr. Merrell provided me with personal encouragement, solid advice, and technical support. In addition, I wish to thank Dr. Susan Crowley and Dr. Patricia Truhn for their thoughtful and insightful feedback provided during the course of the study. Furthermore, I want to extend my gratitude to Dr. Cher King and Mr. Dale Thompsen for their assistance in helping to make this a viable investigation. I would also like to thank the teachers and the staff in the elementary school, whose assistance in the data collection procedures was invaluable. The teachers and staff included Ms. Aiken, Ms. Bowman, Ms. Chadwick, Ms. Dallinga, Ms. Fadel, Ms. Giles, Ms. Hedgepeth, Ms. Malan, Mr. Mann, Ms. Marriott, Ms. Moore, Ms. Tanner, and Mr. Underwood.

Kurt David Michael

CONTENTS

	Page
ABSTRACT	iii
DEDICATION	v
ACKNOWLEDGMENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER	
I. INTRODUCTION AND PROBLEM STATEMENT	1
II. REVIEW OF THE LITERATURE	5
Overview of the Internalizing Disorders Literature	6
Major Subcomponents of Internalizing Disorders	8
Comorbidity Among Internalizing Disorders	13
Implications and Risk Factors of Internalizing Disorders	17
Assessment of Internalizing Disorders in Children	19
Self-Report Instruments	27
Temporal Stability of Internalizing Disorders	33
Summary	41
III. PURPOSE AND OBJECTIVES	43
IV. METHODOLOGY	45
Participants	45
Instrumentation	45
Procedure	50
V. RESULTS	53
Descriptive Statistics	53
Test-Retest Reliability of the ISSC	54
Analysis of Variance with Repeated Measures	56
Proportion and Temporal Stability of "At-Risk" Cases	57

VI. DISCUSSION	61
Test-Retest Reliability of the ISSC	61
Reliability of the ISSC Over Repeated Administrations	63
Temporal Stability of Self-Reported Internalizing Symptoms in the "At-Risk" Group	65
Limitations and Future Research	66
Clinical Implications of Findings	67
REFERENCES	69
APPENDIXES	88
Appendix A: Statement of the Principal Investigator to the Institutional Review Board for Proposed Research Involving Human Subjects	89
Appendix B: Institutional Review Board Letters of Approval	91
Appendix C: Permission Letter from the Ogden City School District	94
Appendix D: Parent Letter and Informed Consent Form	96
Appendix E: ISSC Protocol	98

LIST OF TABLES

Table		Page
1	Test-Retest Reliability Coefficients for the Various Self-Report Instruments	34
2	Sample ISSC Items, Listed by Factor	49
3	Descriptive Statistics for the Total ISSC Scores of 131 Elementary-Age Children for Each Interval	54
4	Means of the ISSC Scores Over Repeated Administrations Based on Gender with t -Test Results and Significance Levels	54
5	Test-Retest Reliability Coefficients for the ISSC Total Scores of 131 Elementary-Age Children	55
6	Test-Retest Reliability Coefficients for ISSC Factor 1 Scores (Negative Affect/General Distress) of 131 Elementary-Age Children	56
7	Test-Retest Reliability Coefficients for ISSC Factor 2 Scores (Positive Affect) of 131 Elementary-Age Children	56
8	Analysis of Variance with Repeated Measures	57
9	“At-Risk” Cases in the Sample as Determined by Total ISSC Scores	59
10	Level of Symptom Endorsement for “At-Risk” Cases for Each Interval	60

LIST OF FIGURES

Figure		Page
1	DSM-IV symptom overlap: Major Depressive Episode and Generalized Anxiety Disorder	14
2	Attenuation of mean ISSC scores across intervals after initial administration	64

CHAPTER I

INTRODUCTION AND PROBLEM STATEMENT

Recent attempts to create empirically sound taxonomies of child psychopathology have yielded two broad dimensions of emotional and behavioral problems, internalizing and externalizing disorders, respectively. Internalizing disorders are a constellation of inner-directed or overcontrolled expressions of distress, whereas externalizing disorders are defined as outer-directed or undercontrolled behavioral problems such as aggression, impulsivity, hyperactivity, and delinquency (Reynolds, 1990). Internalizing disorders encompass a wide variety of symptoms, including depression, anxiety, social withdrawal, and somatic complaints. It has been suggested that the existence of internalizing disorders in children may have negative effects upon their self-esteem, academic achievement, physical health, and future adjustment (Merrell, 1994; Reynolds, 1992a).

Over the past two decades, a great deal of research has been devoted to the understanding of internalizing disorders in children. Prior to the current burgeoning interest in internalizing disorders, the majority of empirical and clinical investigations in child psychopathology focused on externalizing disorders (Reynolds, 1990). This shift in focus has been attributed to several factors. First, because internalizing disorders are generally considered to be insidious and difficult to detect, inquiries into their nature and etiology have likely taken a back seat to more readily observable externalizing behavioral disorders (Reynolds, 1990). This uneven focus would be analogous to the adage that "the squeaky wheel gets the oil." Second, after the American Psychiatric Association (1980, 1987) revised the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-III,

DSM-III-R) during the 1980s, many psychologists became more cognizant of emotional problems in children and adolescents above and beyond those relevant to academic achievement or externalizing behavioral disorders (Reynolds, 1992a). Finally, Reynolds (1992a) suggested that because internalizing disorders typically result in significant distress, misery, and negative outcomes in young people, clinicians and researchers have responded to help understand and assuage their concerns.

The efficacious treatment of virtually every psychopathological disorder is contingent upon an accurate assessment of the pathognomonic symptoms of that disorder (Achenbach, 1985). Traditionally, the evaluation of childhood disorders has relied upon the verbal or written reports of parents, teachers, and other significant figures in the child's environment. However, because internalizing disorders are, in great measure, subjective perceptions of internal distress, they are often not readily or reliably identified by external observers. Outside observers often underestimate the intensity and breadth of a child's emotional experience (Kurdek & Berg, 1987). As a result, several authors have stressed the importance of eliciting the child's perspective through self-report assessment (Flanery, 1990). Subsequently, several self-report measures of internalizing constructs have been developed. Unfortunately, none of these instruments is designed to measure the comprehensive domain of internalizing disorders in children below the age of 11, even though it has been established that children as young as 8 are able to give reliable self-reports (La Greca, 1990; Stone & Lemanek, 1990). This apparent dearth of broad-based instruments for middle- to late-elementary school children creates problems for the assessment of internalizing problems because the various internalizing syndromes often

coexist with one another, therefore limiting the utility of a single-syndrome instrument. (Reynolds, 1992a). For example, depression and anxiety co-occur frequently in the same child, thus the problem of *comorbidity* illustrates the need for a general self-report instrument that measures the broad dimension of internalizing constructs in children (Costello, 1986; Merrell & Walters, 1996).

The newly developed Internalizing Symptoms Scale for Children (ISSC) is a 48-item self-report instrument designed to measure the broad range of internalizing problems in children. The research prototype of the ISSC has been administered to a normative sample of over 2,200 subjects. Preliminary reliability data support the internal consistency of the instrument items and there is evidence of construct validity as indicated by the instrument's sensitivity to various group differences (Merrell & Dobmeyer, 1996; Merrell, Gill, McFarland, & McFarland, 1996; Sanders, 1996).

Additional empirical support (i.e., reliability and validity data) for the ISSC is needed to establish its credibility as an assessment instrument for internalizing symptoms in children. Furthermore, little is known about the temporal stability of self-reported internalizing symptoms in children from a normal population. Various researchers have characterized internalizing disorders as transient when compared to the relative stability of externalizing problems (Fisher, Hasazi, & Cummings, 1984; Graham & Rutter, 1973; McGee et al., 1985), most notably conduct disorder (Offord et al., 1992). However, several recent longitudinal studies have provided evidence to support the notion that internalizing disorders in children from a variety of clinical populations may be relatively stable over time (Cantwell & Baker, 1989; McGee & Williams, 1988; Nolen-Hoeksema,

Girgus, & Seligman, 1992). Additionally, DuBois, Felner, Bartels, and Silverman (1995) provided evidence that self-reported depressive symptoms in a community sample of 435 school-age children were reasonably stable over a period of 2 years. The aforementioned results are promising; however, additional empirical data are needed to better understand the temporal stability of self-reported internalizing symptoms in children from a normal population over short- to medium-length time intervals. Thus, the purpose of this investigation was to gauge the test-retest reliability of the ISSC at several time intervals to provide additional reliability evidence for this instrument and to further the empirical base of knowledge regarding the temporal stability of self-reported internalizing symptoms in children.

CHAPTER II

REVIEW OF THE LITERATURE

Several self-report measures can be used to assess specific constructs within the realm of children's internalizing disorders. The most prominent of these measures are the Children's Depression Inventory (CDI; Kovacs, 1992), the Reynolds Child Depression Scale (RCDS; Reynolds, 1989), the Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985), the State-Trait Anxiety Inventory for Children (STAIC; Spielberger, 1973), and the Youth Self Report (YSR; Achenbach, 1991). With the exception of the YSR, these instruments are syndrome-specific and they are not designed to measure the broad domain of internalizing symptoms. While the YSR purports to measure the breadth of the internalizing domain, it only extends down to age 11, despite evidence that children as young as 8 are capable of giving reliable self-reports (Stone & Lemanek, 1990). This literature review was conducted to establish support for the need to develop a valid and reliable self-report instrument that accurately assesses a broad range of internalizing problems in middle- to late-elementary school-age children. The focus of this inquiry specifically addressed the issues of test-retest reliability and the temporal stability of self-reported internalizing symptoms in children between 8-12 years of age.

A definition and general overview of internalizing disorders is presented, followed by a review of the prevalence, major subcomponents, comorbidity, and risk factors of internalizing problems. A discussion of the importance of using self-report

measures as contrasted with other forms of assessment (i.e., behavior rating scales, direct observation, clinical interviews, sociometric procedures) is provided. In addition, descriptions of current self-report measures designed to assess internalizing constructs are included. Finally, discussions regarding test-retest reliability and the temporal stability of internalizing disorders in children are presented.

Overview of the Internalizing Disorders Literature

In an effort to create empirically based taxonomies of child psychopathology, several authors have categorized emotional and behavioral disorders into two broad-band dimensions of internalizing and externalizing disorders, respectively (Achenbach, 1966, 1985; Achenbach & McConaughy, 1992; Cicchetti & Toth, 1991). Internalizing disorders have been broadly defined as inner-directed or overcontrolled problems (Achenbach & Edelbrock, 1978; Reynolds, 1990). In contrast, externalizing disorders have been described as outer-directed and undercontrolled problems such as aggression, impulsivity, hyperactivity, delinquency, and other overt behavioral problems (Reynolds, 1990). Angold and Costello (1993) argued that “the broad distinction between the emotional (internalizing) disorders and behavioral (externalizing) disorders has stood the tests of time and repeated investigation” (p. 1787).

Major subcomponents of internalizing disorders include depression, anxiety, social withdrawal, and somatic complaints (Merrell, 1994; Reynolds, 1992a). These problems have been found to be interrelated clinically and they have been shown to be strongly associated in factor-analytic studies (Ollendick & King, 1994). Prevalence rates

of internalizing disorders in children vary depending upon the particular disorder under investigation and the diagnostic criteria being used. However, prevalence estimates for particular childhood internalizing disorders have ranged from 2.0% for depression to 8.9% for anxiety in normal samples (Anderson, Williams, McGee, & Silva, 1987; Costello, 1989).

Despite the fact that the broad-band dimension of internalizing disorders has been empirically supported (Achenbach & Edelbrock, 1989; Quay & La Greca, 1986), a thorough understanding of the entire spectrum of internalizing symptoms and how they are interrelated has been inhibited by several problems. Because internalizing disorders include internal or subjective perceptions, an accurate assessment of such disorders through the use of self-report measures may be hindered by developmental problems (Clarizio, 1984), limited self-understanding (Stone & Lemanek, 1990), lack of emotional insight (La Greca, 1990), and reading level (Prout & Chizik, 1988). Furthermore, Costello (1986) reported that “it is by no means certain that the younger school-age child can recognize the sustained unhappiness in [internalizing disorders such as] depression” (p. 565). However, despite these limitations, several authors have stressed the importance of self-report when evaluating internalizing symptoms in school-age children (Finch, Saylor, Edward & McIntosh, 1987; Flanery, 1990; La Greca, 1990; Merrell, 1994; Saylor et al., 1984).

Other methods of assessing internalizing disorders have yielded inconsistent results. For example, direct behavioral observations and behavioral checklists often yield discrepancies among child, parent, and teacher observations and reports (Achenbach,

McConaughy, & Howell, 1987; Reynolds, 1992a). Another problem that obscures the understanding of internalizing disorders is the issue of *comorbidity*, or the co-occurrence of two or more disorders in the same child. Despite the evidence that several narrow-band internalizing disorders often occur together (e.g., depression and anxiety), there is a great deal of variation and overlap in symptom presentation in children with internalizing disorders (Ollendick & King, 1994). Subtle distinctions between the various internalizing symptoms are often difficult to make. Subsequently, various authors have suggested that an appropriate assessment instrument for internalizing symptoms should be broad enough to accurately identify several different constellations of internalizing symptoms (Achenbach, 1985; Costello, 1986; Reynolds, 1992a).

Major Subcomponents of Internalizing Disorders

After the broad dimension of internalizing problems was identified and empirically supported (Achenbach, 1985; Achenbach & Edelbrock, 1978; Cicchetti & Toth, 1991), several authors conducted multivariate analyses to identify the major subcomponents of internalizing disorders. In a factor analysis of the behavior problem items on the Child Behavior Checklist (CBCL) for a large sample of children referred for mental health treatment, Achenbach and Edelbrock (1983) found that the internalizing dimension contained several factors, including depressed, anxious, somatic complaints, social withdrawal, schizoid, immature, and obsessive-compulsive. The factors of depressed, anxious/schizoid, social withdrawal, and somatic complaints were consistent across gender. These findings have been replicated in other studies by different authors

(Achenbach, Conners, Quay, Verhulst, & Howell, 1989; Achenbach & McConaughy, 1992). It is important to note, however, that the internalizing-externalizing distinction is not a perfect dichotomy and subsequent attempts to break down each dimension into smaller parts may be difficult (Ollendick & King, 1994).

General definitions for the four major subcomponents of internalizing disorders (i.e., depression, anxiety, social withdrawal, and somatic complaints) are presented in this section. Prevalence rates for each of the subcomponents are given. However, due to the differences in instrumentation, sampling techniques, diagnostic procedures, and population samples, there is a great deal of variation in the prevalence estimates (Rutter, 1989).

Depression

Prior conceptualizations of childhood depression were presumed to be distinguishable from adult forms of depression. Some of the conceptualizations included depressive equivalents or “masked depression,” which were purportedly manifested by overt behavioral problems such as delinquency, hyperactivity, and aggression (Cytryn & McKnew, 1972; Glaser, 1967). However, based on current research and clinical opinion, depression in children, for the most part, is characterized and identified in many of the same ways as depression in adults (Puig-Antich, 1982). Some authors have argued that while the differences in distinguishing childhood depression from the adult forms of the disorder are minor, developmental factors (e.g., language, cognitive abilities, emotional insight) must be taken into account when attempting to identify and classify childhood

depression (Carlson & Garber, 1986).

As a symptom, depression is characterized by a dysphoric or unhappy mood state. Defined as a syndrome, depression consists of a constellation of behavioral and emotional symptoms that do not simultaneously exist by chance (Rehm & Tyndall, 1993). For example, when a dysphoric mood is combined with labored psychomotor functioning, cognitive difficulties, and a lack of motivation, these symptoms, if experienced simultaneously, are often construed as evidence of a depressive syndrome. Establishing the existence of a depressive *disorder* depends largely upon how long the depressive syndrome has persisted. In the *Diagnostic and Statistical Manual of Mental Disorders*, (4th ed.; APA, 1994), a diagnosis of either a depressive episode or a depressive disorder depends upon whether a certain number of criteria have been met. The criteria are made up of several emotional, cognitive, and behavioral symptoms, including dysphoric mood, anhedonia, impaired academic, interpersonal, and social functioning, difficulty concentrating, sleep and appetite disturbance, and fatigue.

Estimates of the prevalence of childhood depression are varied, ranging from 2% to 17.9% (Kanshani et al., 1983; Lefkowitz & Tesiny, 1985; Silver, 1988). In a critical evaluation of epidemiological studies of childhood depression, Fleming and Offord (1990) suggested that the variation in prevalence rates is due primarily to methodological flaws in the research designs (e.g., sampling bias, small samples, inconsistent measurements and diagnostic procedures, etc.). Nevertheless, even the lowest estimates are high enough to create cause for concern. Despite the variation in these percentages, Reynolds (1990) suggested that these figures underestimate the actual prevalence of

childhood depression. Reynolds (1992b) reported that approximately one out of six youngsters receiving psychiatric services have been formally diagnosed with an affective disorder, thus making childhood depression “one of the most prevalent and pervasive forms of psychopathology in this age group” (p. 150).

Anxiety

Anxiety is defined as a tense emotional state characterized by feelings of distress, fear, physiological arousal, and maladaptive patterns of thinking and behavior (Strauss, 1990). In the *DSM-III-R* (APA, 1987), childhood anxiety disorders were classified into three subtypes, including Separation Anxiety Disorder, Overanxious Disorder, and Avoidant Disorder. Each subtype of anxiety disorder is said to have distinguishing features. Separation Anxiety Disorder is characterized by “distress about separation from home or from a major attachment figure” (Strauss, 1990, p. 142). The essential feature of Overanxious Disorder is excessive or unrealistic worry about the future. Avoidant Disorder is a condition whereby the child demonstrates excessive fearfulness and avoidance of social situations to the point where social functioning and peer relationships are significantly impaired. However, in the *DSM-IV* (APA, 1994), only Separation Anxiety Disorder retained its previous classification status. Overanxious Disorder and Avoidant Disorder were subsumed under Generalized Anxiety Disorder and Phobias, respectively. Some of the *DSM-IV* criteria for anxiety disorders include restlessness, fear, distress, difficulty concentrating, physical arousal, irritability, sleep disturbance, muscle tension, and patterns of behavioral avoidance.

Despite its somewhat broad and diffuse definition, it has been suggested that childhood anxiety disorders are one of the most prevalent categories of child and adolescent disorders (Bernstein & Borchardt, 1991). Prevalence estimates for childhood anxiety disorders have ranged from 3.5% (Anderson et al., 1987) to 8.9% (Costello, 1989).

Social Withdrawal

Social withdrawal is characterized by a reluctance to engage in social situations, excessive fear of unfamiliar stimuli, and behavioral withdrawal and isolation (Kauffman, 1989). Social withdrawal is considered to be one of the major correlates of anxiety and depression and is frequently cited as a category of behavioral deficits associated with internalizing disorders (Kauffman, 1989). Quay and La Greca (1986) estimated the prevalence rates of severe social withdrawal in children to be approximately 2%.

Somatic Complaints

Werry (1986) defined somatic complaints as a group of disorders characterized by physical symptoms for which there appears to be no physical explanation. Merrell (1994) noted that "somatic symptoms associated with internalizing characteristics are presumably psychological in origin" (p. 190). Common somatic complaints include headaches, abdominal pain, vomiting, and eye problems. Garralda (1992) suggested that because most children are not adept at verbalizing their emotions, they often use somatic complaints as an alternate method of communicating distress. Greene and Thompson (1984) estimated that between 15% and 20% of school children present with somatic

complaints, 90% of which have no known physical cause.

Comorbidity Among Internalizing Disorders

According to Stedman's Medical Dictionary (1995), *comorbidity* is defined as "a concomitant but unrelated pathological or disease process" (p. 174). For example, an individual might be suffering from both lung cancer and Hepatitis B at the same time and thus be considered to have comorbid medical illnesses. These two disease processes are essentially independent of one another, with different etiologies, symptom presentations, and progression patterns.

While the medical definition of comorbidity implies that illnesses are concomitant but unrelated, the use of the term "comorbidity" in the psychological and psychiatric literature is less well-defined. Unlike many medical illnesses, psychological disturbances are not discrete illnesses and are therefore more difficult to assess, diagnose, and classify, due, in part, to the overlapping nature of the various symptom clusters (Adams & Cassidy, 1993). For example, two internalizing disorders in the *DSM-IV* (1994), Major Depressive Episode and Generalized Anxiety Disorder, have overlapping diagnostic criteria, including irritability, difficulty concentrating, sleep disturbance, and fatigue (see Figure 1). Consequently, individuals who present with these symptoms would meet some of the diagnostic criteria for Major Depressive Episode (MDE) while simultaneously satisfying some of the diagnostic parameters of Generalized Anxiety Disorder (GAD), thus making an accurate differentiation between the two diagnoses problematic.

Aided by advances in microbiology, genetics, and biochemistry as well as

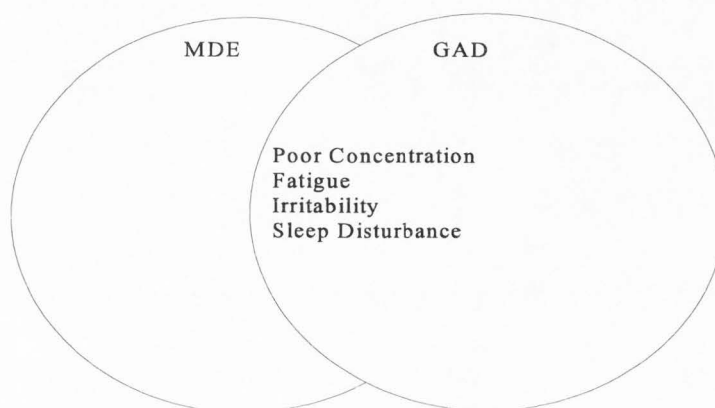


Figure 1. DSM-IV (APA, 1994) symptom overlap: Major Depressive Episode and Generalized Anxiety Disorder.

advanced diagnostic procedures, many medical diseases can be reliably identified and diagnosed (Garfield, 1993). By comparison, current assessment and classification procedures for psychopathology are relatively unreliable (Garfield, 1993). Nevertheless, researchers and clinicians have continued the quest to understand the complex relationships between various psychological disorders.

Determining whether two psychological disturbances are comorbid in the same individual at the same time depends on various considerations. One must consider whether the comorbidity is the result of measurement error, similarity in self-report tendencies, diagnostic imprecision, or the actual diagnostic criteria. Angold and Costello

(1993) stated that researchers and clinicians must evaluate whether patterns of comorbidity are “artifacts of the methods of data collection, data aggregation for diagnostic purposes, or the nosology itself” (p. 1786).

In spite of these diagnostic limitations, there is mounting evidence that the existence of more than one psychological disorder in the same individual is prevalent and can lead to poor outcomes. In a recent longitudinal study of over 1,000 children who were followed from birth to age 21, Newman and colleagues (1996) reported that nearly half of the subjects who evidenced a psychiatric disorder during the course of the study also had comorbid diagnoses at the age of 21. In addition, the authors indicated that “comorbidity was associated with severity of impairment” (p. 552).

As previously mentioned, the major subcomponents of internalizing disorders are not discrete categories; thus the co-occurrence or comorbidity of two or more of the subcomponents in the same child is not only possible, it is common. Current comorbidity estimates for depression and anxiety range from 15.9% to 61.9% (Brady & Kendall, 1992). Anderson et al. (1987) examined a nonclinical sample of 63 children and found that 15.9% qualified for both an anxiety disorder and a depressive disorder. However, Costello and colleagues (1988) reported much lower estimates of comorbidity in a nonclinical group of pediatric primary care patients, with coexisting symptoms of depression and anxiety appearing in 0.8% of the sample. In clinical samples, the comorbidity rates have been much higher. In a sample of hospitalized children, Carey, Finch, and Imm (1989) reported that 55.2% of the sample had diagnosable disorders of both depression and anxiety. In a group of outpatient children and adolescents who

presented with anxiety disorders, Strauss, Last, Hersen, and Kazdin (1988) reported that 28.3% of the sample also met criteria for depressive disorders.

Comorbidity estimates between anxiety and somatic complaints have been cited as well. King and Ollendick (1989) reported that somatic complaints are often endorsed in children with school phobias. In a clinical outpatient group of children and adolescents, Last (1991) found that 60% of the sample was comorbid for anxiety and somatic complaints.

In summary, the comorbidity rates between some of the major subcomponents of internalizing disorders are varied. The variation in comorbidity estimates has been attributed to “rather crude diagnostic criteria” (Angold & Costello, 1993, p. 1786), unreliable data collection techniques (Garfield, 1993), similarity in self-report rather than construct overlap (Norvell, Brophy, & Finch, 1985), and the fact that many internalizing disorders may be clinically related (Ollendick & King, 1994). Despite the discrepant findings, several authors have suggested that the overall comorbidity rates are large enough to be considered clinically meaningful (Kendall, Kortlander, Chansky, & Brady, 1992; Newman et al., 1996; Reynolds, 1992a). In light of the variation in comorbidity estimates and the diverse presentation of internalizing symptoms, various researchers have recommended assessment practices that emphasize broad-band instruments and the solicitation of information from several sources (Achenbach et al., 1987; Finch et al., 1987; Kazdin, 1988; Reynolds, 1992c).

Implications and Risk Factors of Internalizing Disorders

Persistence and Long-Term Implications

Although Quay and Werry (1986) suggested that certain internalizing disorders may not carry the “foreboding prognosis” that is often associated with conduct (externalizing) disorders, other researchers have asserted that internalizing disorders may lead to long-term negative outcomes (Cantwell, 1990; Reynolds, 1992a). For example, Fischer et al. (1984) followed a sample of preschool children identified as having internalizing and externalizing problems over several years. The authors reported positive correlations between preschool internalizing behaviors and similar problems in late elementary school. In addition, Fischer and colleagues (1984) found an inverse relationship between preschool internalizing behaviors and later social competence. Anxiety disorders in childhood also appear to persist into adolescence. In a follow-up study of children who were diagnosed with anxiety disorders, Cantwell and Baker (1989) found that approximately 50% of the children presented with a substantial number of anxious symptoms in early adolescence.

Kovacs (1985) reported that childhood depression may be a precursor or risk factor for psychological disorders in adulthood. In a longitudinal study of depressive disorders in children, a significant number of youngsters remained symptomatic for 5 years or more, even when treatment was implemented (Kovacs et al., 1984). These findings appear to contradict the popular belief that young children only suffer from brief and episodic depressive disorders.

Poor Self-Concept

Self-concept is broadly defined by how a person self-evaluates several areas of functioning, including academic competence, personal appearance, and social dexterity. This broad definition of self-concept has been used to illustrate the *multidimensional* nature of the construct. Harter (1990) suggested that each aspect of a child's self-concept is relatively independent. For example, if a child feels academically competent, but also feels quite insecure about his/her physical appearance, it does not necessarily mean that one type of self-concept will or will not take precedence over the other in terms of the child's global self-concept. Merrell (1994) suggested that a person's self-evaluative tendencies have important implications for internalizing disorders. There has been evidence to suggest that a poor self-concept is associated with depression (Kazdin, 1988), somatic complaints (Walker & Greene, 1989), and impaired academic achievement (Bloom, 1976).

Impaired Academic Performance

Various researchers have discovered a relationship between internalizing disorders such as depression, anxiety, social withdrawal, and somatic complaints and poor academic performance (Appolloni & Cooke, 1977; Last, 1991) as well as school dropouts (Fleming & Offord, 1990). Quay and La Greca (1986) reported that highly anxious children perform more poorly on measures of academic achievement when compared to less anxious peers. King and Ollendick (1989) suggested that children with anxiety and school phobias experience levels of distress that hinder academic and social development.

Risk of Suicide

One internalizing disorder in particular, namely depression, has been linked to suicide (Bettes & Walker, 1986; Rao, Weissman, Martin, & Hammond, 1993). Kovacs, Goldston, and Gatsonis (1993) reported that Major Depressive Disorder and Dysthymic Disorders were associated with significantly higher rates of suicide than were Adjustment Disorder with Depressed Mood and nondepressive disorders in a mixed sample of children between the ages of 8 and 13. While a clear relationship between childhood depression and suicide has not been established, Smith (1992) described the nature of suicidal behavior in children as an internalizing disorder. However, Reynolds (1992b) cautioned that a significant number of depressed youngsters are not necessarily at risk for suicide and, conversely, a number of youngsters who exhibit suicidal behaviors are not depressed. Nonetheless, hopelessness, coupled with depression, increases the probability of suicide attempts (Smith, 1992). In addition, Kovacs and colleagues (1993) found that “in the presence of affective disorders, comorbid conduct and/or substance abuse disorders further increased the risk of suicide attempts” (p. 8).

Assessment of Internalizing Disorders in Children

Widely endorsed methods of childhood assessment often include several basic tenets. La Greca (1990) suggested that a comprehensive child assessment must include: multiple evaluators in the child’s environment, multiple methods of data collection, and the solicitation of the child’s perspective. La Greca (1990) emphasized the importance of a multimethod-multisource procedure that takes into account, “the limitations inherent in

any one procedure” (p. 8).

The primary methods for assessing internalizing disorders are direct behavioral observation, behavior rating scales, sociometric approaches, clinical interviews, and self-report measures. This assessment information may be solicited from a variety of sources (i.e., parents, teachers, children, peers).

Behavioral Observation

Direct behavioral observation may provide important information about internalizing disorders in children because several internalizing problems have behavioral manifestations. For example, Kazdin (1988) asserted that observable behaviors such as decreased motor activity, labored speech, and limited social contact are symptoms of depression. However, because many internalizing disorders (e.g., depression) are inner-experienced subjective states of distress, assessing this phenomenon may be difficult (Reynolds, 1992c). Consequently, children with internalizing problems may not be easily identified by parents and teachers through behavioral observation.

Behavioral Rating Scales

The use of behavioral rating scales has proven to be an integral part of a multimethod child assessment (La Greca, 1990). Edelbrock (1983) suggested that the use of behavioral rating scales such as the Child Behavior Checklist (CBCL) and the CBCL Teacher Report Form (Achenbach & Edelbrock, 1983) are efficient and cost-effective ways of obtaining data on child behavior. Behavior rating scales often solicit information from parents (or primary care givers) and teachers. Achenbach (1991) asserted that

parents and teachers are the first and second most important sources of information about a child's competencies and concerns, respectively. However, behavior rating scales have limitations, especially with respect to the accurate assessment of internalizing disorders in children. This limitation of behavior rating scales is due primarily to the insidious nature of internalizing disorders and the subsequent difficulty in measuring or evaluating them. Another limitation of using behavioral rating scales to assess internalizing disorders is the high level of disagreement amongst informants (Achenbach et al., 1987; Kazdin, 1989; Reynolds & Graves, 1989). In general, parents typically underreport affective disorders, anxiety disorders, and somatic complaints of their children (Weissman et al., 1987). These low rates of agreement underscore the importance of using alternate and/or additional methods of assessing internalizing disorders in children.

Sociometric Procedures

Sociometric approaches to child assessment emphasize peer report. Sociometric approaches may thus be used to gather data about the observable or perceived characteristics of internalizing disorders. The Peer Nomination Inventory for Depression (PNID; Lefkowitz & Tesiny, 1980) is the most widely used sociometric measure of depression in children (Merrell, 1994). However, as previously noted, internalizing disorders are predominantly subjective internal states and are not readily detected or observed by external informants. Furthermore, young children may not be able to identify subtle differences or characteristics of mood states in others (Merrell, 1994).

Self-Report

Because self-report assessment is a primary focus of this study, more detail will be provided on this form of assessment than on other forms. Self-report data, whether obtained through structured/unstructured clinical interviews or paper-and-pencil measures, play a vital part in the accurate assessment of internalizing disorders in children (Flanery, 1990; La Greca, 1990; Martin, 1988). However, Merrell (1994) noted that with children “there is often a reluctance to relinquish the use of external methods of assessment, due to the supposedly questionable accuracy of information obtained through self-report methods” (p. 194). La Greca (1990) reaffirmed the importance of the child’s perspective, especially in the assessment of internalizing disorders and characterized the child’s subjective evaluation of internal distress as “paramount.” External evaluations of a child’s internal state are often inaccurate, unreliable, and subject to significant observer bias (Edelbrock, Costello, Dulcan, Conover, & Kalas, 1986; La Greca, 1990).

Interviews

Clinical interviews are one of the most commonly used methods of assessing internalizing symptoms in children (Angold & Costello, 1993; Miller, Boyer, & Rodoletz, 1990). Edelbrock, Costello, Dulcan, and Conover (1985) characterized the clinical interview as “the cornerstone of child clinical assessment” (p. 265). A number of clinicians and researchers have chosen to follow a structured or semistructured format when soliciting self-report information from children because they have been found to yield more reliable and comprehensive data than “free-form” interviews (Edelbrock et al.,

1985). Some of the more prominent interview formats include: the Schedule for Affective Disorders and Schizophrenia for School-Aged Children (K-SADS; Puig-Antich & Chambers, 1978), the Diagnostic Interview Schedule for Children (DISC; Costello, Edelbrock, Dulcan, Kalas, & Klaric, 1984), the Diagnostic Interview for Children and Adolescents-Revised (DICA-R; Reich & Welner, 1988), and the Child Assessment Scale (CAS; Hodges, Kline, Fitch, McKnew, & Cytryn, 1981). Many of these instruments have been developed in accordance with current diagnostic systems (e.g., DSM) and they are often designed to support or rule out particular psychiatric diagnosis.

Structured and semistructured interviews provide the clinician with a rich opportunity to gather important information in a flexible manner. However, this supposed strength of using structured/semistructured clinical interviews can quickly sour into a pronounced liability. Finn and Kamphuis (1995) lamented the fact that virtually all types of clinical interviews are unreliable and subject to the personal biases of the clinician. In other words, even when the interview is fairly well-structured, evaluators “see what they expect to see” and continue to probe in areas in which they expect to find problems (Angold & Costello, 1993). Another limitation of using structured/semistructured interviews in the assessment of internalizing symptoms in young children is that some researchers have questioned a child’s ability to give reliable self-reports over brief (1- to 2-week) time intervals (Costello, 1986). In a study of 242 disturbed children who were interviewed using the DISC, Edelbrock and colleagues (1985) reported that the average 1-week test-retest reliability coefficient for children between the ages of 6-9 (.43) was significantly lower than the average coefficients for the older cohorts (10-13 = .60;

14-18 = .71). Boyle and colleagues (1993) reported similar findings with a revised version of the DICA (DICA-R; Reich & Welner, 1988) in a community sample of children between the ages of 6 and 16. Edelbrock and colleagues (1985) suggested that an age-related increase in the “reliability of the child’s report was expected, given the child’s improving cognitive, memory, and language skills” (p. 273). However, the authors cautioned that simply because young children below the age of 10 appear to give less reliable self-reports, this should not deter clinicians and researchers from using structured interviews since they also help to establish rapport and provide opportunities to observe mental status, motor behavior, and verbosity.

In a recent study using the DISC-R (DISC-R; Schaffer, Schwab-Stone, & Fisher, 1993), Schwab-Stone and colleagues (1993) interviewed a clinical sample of 74 children between 11 and 17 years old twice over a time period ranging from 1 to 3 weeks. Test-retest reliabilities were reported for childhood disorders, including internalizing diagnoses such as Major Depressive Episode (.77) and Separation Anxiety Disorder (.72). In study of a sample of 375 normal and referred children between the ages of 9-17 utilizing still another version of the DISC (Version 2.1), Jensen and colleagues (1995) reported that the test-retest reliability coefficients of the revised DISC were “consistent or superior to those reported in previous studies” (p. 61). However, it was noted that closely spaced or repeated DISC interviews resulted in significant diagnostic attenuation on retest (Jensen et al., 1995). The authors suggested that the test-retest attenuation phenomena might have been due to a decrease in the self-reported symptomology below the diagnostic threshold, inconsistent interview and diagnostic procedures, and regression to the mean.

Another limitation of structured/semistructured interviews worthy of mention is the fact that they are labor-intensive and expensive to administer, score, and interpret (Merrell, 1994). Nonetheless, structured/semistructured interviews remain a valuable tool in the assessment of internalizing disorders in children.

Objective Self-Report

Whereas structured or semistructured interviews are often used to determine whether patients' self-reported symptoms reach a "diagnostic threshold," objective self-report instruments are typically used to assess the degree to which respondents endorse clinically significant symptomology relevant to a particular problem area (La Greca, 1990). Several excellent objective self-report instruments have been designed to assess specific internalizing problems in children. The Children's Depression Inventory (CDI; Kovacs, 1992), the Reynolds Child Depression Scale (RCDS; Reynolds, 1989), the Revised Children's Manifest Anxiety Scale (RCMAS; Reynolds & Richmond, 1985), and the State-Trait Anxiety Inventory for Children (STAIC; Spielberger, 1973) have been widely used to assess subcomponents of internalizing disorders. In addition, the Youth Self Report (YSR; Achenbach, 1991) is designed to measure common internalizing constructs such as depression and anxiety as well as other internalizing subcomponents such as withdrawal and somatic complaints. While the YSR may be one of the few objective self-report instruments that purports to measure somatic complaints and withdrawal in a relatively independent manner, many of the so-called single syndrome instruments (depression, anxiety) contain items that are correlated with these internalizing

subcomponents (Merrell, 1994). For example, Reynolds (1989) reported that “the RCDS measures a range of symptomology associated with depression including cognitive, motoric-vegetative, somatic, and interpersonal symptoms” (p. 1).

These instruments tend to utilize a common response format, in that the child subject is presented with a series of statements regarding the presence or absence of specific symptoms, and then rates how true these statements are for them, or how often they occur. The authors of most objective self-report assessment devices have established cutoff scores to operationalize clinically relevant levels of symptomology for their instruments, that is, criterion-related validity (Reynolds, 1989; Flanery, 1990). A number of metrics have been used to indicate clinical cutoff points for various self-report instruments, including raw scores, T-scores, and percentile ranks. A response set that is 1.5 to 2.0 standard deviations (SDs) above the mean on self-report inventories (MMPI, CDI, etc.) has generally been considered to be a good indication of clinically relevant self-reported symptomology, assuming that certain assumptions regarding the sample (normally distributed, clinical vs. nonclinical, random sample) have been considered (Merrell, 1994). While the establishment of clinical cutoff scores (e.g., 1.5 to 2.0 SDs above the mean, T-scores of 65 - 70) on self-report measures is not equivalent to a formal diagnosis (Reynolds, 1989), it often provides valuable information regarding the severity of certain symptomology (i.e., criterion) in comparison to a particular reference group as long as the base rates for the criterion (e.g., depression, anxiety) are taken into account (Finn & Kamphuis, 1995).

Despite the fact that objective self-report measures are important in the

assessment of internalizing disorders in children, these instruments have limitations as well. Anastasi (1988) cautioned that the majority of items on self-report inventories “have one answer that is recognizable as socially more desirable or acceptable than the others” (p. 549). As a result, these types of biased response styles (i.e., social desirability, faking good, faking bad) may contribute to the error variance of the instrument (Anastasi, 1988; Borg & Gall, 1989). Other potential limitations of self-report measures include: a child’s ability to understand and report their emotions (Clarizio, 1984; La Greca, 1990), a child’s ability to reliably report subjective states of internal distress (Edelbrock et al., 1985), and whether the instrument has an age-appropriate reading level (Prout & Chizik, 1988). Nonetheless, an accurate assessment of internalizing symptoms in children should include self-report data as part of a multimethod, multisource evaluation (Achenbach et al., 1987; Finch et al., 1987; Flanery, 1990; Kazdin & Petti, 1982; La Greca, 1990; Merrell, 1994; Saylor et al., 1984). Each of the aforementioned objective self-report instruments will be reviewed in the following section.

Self-Report Instruments

Children’s Depression Inventory

The CDI is a 27-item self-report instrument suitable for school-aged children and adolescents between the ages of 7 and 17. It was originally developed in the late 1970s in response to the need for a self-report instrument for depression in children. The CDI quantifies a range of depressive symptomology endorsed, including dysphoric mood,

anhedonia, interpersonal behavior, vegetative tendencies, and self-evaluation.

Respondents are asked to select statements that best describe their feelings during the past 2 weeks. Each of the 27 items consists of three choices (0, 1, 2) with the higher scores indicating increased severity. The total score can range from 0 to 54. Administration of the CDI is relatively simple and usually takes less than 30 minutes to complete the entire process. The determination of clinically significant depressive symptomology is based upon T -scores and general cutoff points for various groups depending on gender, age, and whether the subject is from a clinical or normal sample. Kovacs (1992) suggested that when administered in groups of children not expected to have problems, a T -score of 70 (i.e., 2 standard deviations above the mean) is generally a good indicator of clinically significant depressive symptoms. However, when the CDI is administered for screening purposes, Kovacs (1992) recommended using lower cutoff scores to minimize the possibility of false negatives.

The CDI has been widely used and extensively researched. Normative data were based upon a sample of 1,463 Florida school children (grades 2-8) in the mid-1980s. The psychometric properties of the CDI have been reported in the manual and in a number of studies. Internal consistency coefficients as reported in the manual have ranged from .71 to .89. In a normal sample of children between the ages of 7 and 12, Finch et al. (1987) reported 2-week, 4-week, and 6-week test-retest reliability coefficients of .82, .66, and .67, respectively. In addition, Weiss et al. (1991) reported a 4-month test-retest reliability coefficient of .54. Concurrent validity of CDI has been examined by correlating CDI scores with other measures of internalizing constructs such as the RCMAS (.65; Kovacs,

1992), the RCDS (.68-.79; Reynolds, 1992b), and the STAIC (.71 State, .81 Trait; Smith, Mitchell, McCauley, & Calderon, 1990). The CDI has been found to have significant relationships with all of these measures, an indication of concurrent validity as well as the comorbidity between several internalizing constructs. These findings support the need to develop a broad-based measure of internalizing symptoms.

Reynolds Child Depression Scale (RCDS)

The RCDS is a 30-item self-report instrument that assesses depressive symptomology in children 8-12 years of age. The items were based primarily on the depressive symptoms found in the *DSM-III* (APA, 1980). The items are either endorsed or disavowed based upon a 4-point scale, with higher numbers indicating an increasingly severe endorsement of depressive symptomology. In order to operationalize the clinical threshold of depressive symptomology, the author of the RCDS empirically established clinical cutoff scores based on raw score points (Reynolds, 1989).

The psychometric properties of the RCDS have been reported to range from acceptable to excellent, with internal consistency figures averaging around .90. Two-week test-retest reliability coefficients have been reported at .82 (Breen, 1987) while stability estimates at 4-week intervals have been reported to be .85 (Reynolds & Graves, 1989). Concurrent validity of the RCDS has been based upon correlations with other self-report measures depression such as the CDI (.68 to .79) and measures of anxiety such as the RCMAS and the STAIC (.60 to .67).

While the RCDS remains as a good measure of childhood depression, the

significant overlap that exists between the RCDS and measures of anxiety (RCMAS, STAIC) makes differential assessment problematic.

Revised Children's Manifest Anxiety Scale (RCMAS)

The RCMAS is a 37-item self-report instrument designed to measure trait anxiety or the propensity to be anxious across time and situations. The theoretical underpinnings of the RCMAS are based upon the notion of trait anxiety as described by Taylor (1951) and Spielberger (1972). Trait anxiety has been defined as anxiety that is relatively stable over time and settings, whereas "state" anxiety has been described as anxious symptomology that fluctuates across time and environmental settings (Spielberger, 1972). The instrument is appropriate for children and adolescents who range in age from 6 to 19 years old. Children are asked to respond to items by circling either "yes" or "no." The four subscales on the RCMAS include physiological anxiety, worry/oversensitivity, social concerns/concentration, and a 9-item lie scale. On the RCMAS, the determination of clinically relevant symptomology is based upon scaled scores and percentile ranks in comparison to the normative group.

Internal consistency estimates for the total anxiety score have generally been reported in the low to middle .80s. Test-retest reliability data on the RCMAS indicate that it is stable over 1-week (.88) and 5-week (.77) intervals (Wisniewski, Mulick, Genshaft, & Coury, 1987). Test-retest reliability coefficients based upon the total anxiety score were reported to be .68 at a 9-month interval (Reynolds, 1981). The authors suggested that in light of the lengthy interval (9 months), a test-retest coefficient of this

magnitude supports the stability of chronic (trait) anxiety over time. Convergent validity correlations between the RCMAS and the STAIC Trait scale have been reported to be .78, lending support to the idea that the RCMAS is a good measure of trait anxiety. Divergent validity coefficients between the RCMAS and the STAIC State (acute anxiety) scale have been reported to be extremely low, thus providing additional support for using the RCMAS as a measure of chronic anxiety.

State-Trait Anxiety Inventory for Children (STAIC)

The STAIC is a self-report measure consisting of 20 items designed to assess trait anxiety and 20 items that purport to measure state anxiety (anxious symptomology that vary across time and settings). The STAIC is a downward extension of the State-Trait Anxiety Inventory for adolescents and adults (Spielberger, Gorsuch, & Lushene, 1970) and is appropriate for children who range in age from 9-12 years old. Similar to the CDI, respondents are asked to endorse or disavow symptoms of anxiety based upon a three-point scale (1, 2, 3) with higher scores indicating more severe symptomology. The severity of anxiety symptoms as reported on the STAIC is determined by calculating T-scores and percentile ranks and then comparing these response sets with those of a particular reference group. The state and trait 20-item scales can be administered together or separately and require approximately 10 minutes per scale to administer.

Psychometric properties of the STAIC have been reported to be good, based upon the test manual and the research literature. Internal consistency coefficients for both the state and trait scales have been reported to be in the .80s. Test-retest reliability figures at

6-week intervals range from .65 to .71 for the trait anxiety scale, and .31 to .41 for the state anxiety scale (Spielberger, 1973). Concurrent validity estimates between the STAIC and other measures of internalizing constructs have been reported as adequate.

Youth Self-Report (YSR)

The YSR is a self-report instrument for children between the ages of 11-18 and requires fifth-grade reading skills. The YSR can be administered independently; however, it is often used in concert with the CBCL-Parent Form and the CBCL-Teacher Report Form. The YSR contains 103 statements about various problem behaviors, which the respondent is asked to rate as a “0” (not at all), a “1” (somewhat or sometimes true), or a “2” (very true or often true). The 103 items are scored along the two broad-band dimensions (i.e., internalizing and externalizing), eight narrow-band syndromes, including withdrawal and somatic complaints, and a total score. Scores in each of these problem areas are compared with responses from a normative group of children of the same sex in the same age range. Ratings in one or more problem areas that are higher than 98% of the normative sample are considered to be “clinically significant,” thus warranting further attention.

Psychometric data on the YSR reported in the manual indicate that it is sufficiently correlated with the CBCL-Parent Form and the CBCL-Teacher Report Form (in the .40s). One-week test-retest reliability coefficients were reported to range between .83 and .87 for the broad-band dimensions. Achenbach and Edelbrock (1987) reported that in a sample of 50 nonreferred adolescents, the test-retest reliability coefficients were

.81 over a 1-week period. In the same study, the authors reported a 6-month test-retest reliability of .69 for a group of referred adolescents.

In general, all of the aforementioned self-report instruments that assess internalizing constructs have been widely used and possess adequate to excellent psychometric properties (test-retest reliability coefficients are summarized in Table 1). However, because internalizing problems are often comorbid, instruments designed to measure a unitary internalizing construct are limited in their ability to broadly assess the various internalizing symptom presentations. Reynolds (1992c) argued that there is a need to better understand the relationship amongst the various internalizing problems in children for the purposes of screening, assessment, differential diagnosis, and more precise treatment recommendations. The development of a valid and reliable broad-band measure of internalizing symptoms such as the ISSC would address such a need.

Temporal Stability of Internalizing Disorders

There are two important ways to conceptualize the temporal stability of internalizing disorders in children. First, temporal stability might denote the extent to which children are able to give reliable self-reports of internalizing symptoms over time. Second, temporal stability might refer to how persistent and chronic certain internalizing symptoms and disorders are over time. While these two conceptualizations are similar and interrelated, they are not the same. The reliability of children's self-report speaks to their ability to consistently report subjective internal states over time, whereas estimates of temporal stability purportedly provide evidence as to how stable a particular

Table 1

Test-Retest Reliability Coefficients for the Various Self-Report Instruments

Instrument	Sample	Author(s)	Interval	r
CDI	Normal	Finch et al., 1987	2 weeks	.82
	Normal	Finch et al., 1987	4 weeks	.66
	Referred	Weiss et al., 1991	16 weeks	.54
RCDS	Normal	Breen, 1987	2 weeks	.82
	Normal	Reynolds & Graves, 1989	4 weeks	.85
RCMAS	Normal	Wisniewski et al., 1987	1-5 weeks	.88-.77
	Normal	Reynolds, 1981	9 months	.68
STAIC	Normal-Trait	Spielberger, 1973	6 weeks	.65-.71
	Normal-State	Spielberger, 1973	6 weeks	.31-.41
	Referred	Finch et al., 1984		.44
YSR	Normal	Achenbach & Edelbrock, 1987	1 week	.81
	Referred	Achenbach & Edelbrock, 1987	6 months	.69

psychological construct (e.g., depression) is over time. Both conceptualizations are similar in that the estimates of stability or consistency are often determined by calculating test-retest reliability coefficients. The concepts of test-retest reliability, temporal stability, and the reliability of children's self-report, as well as how they are related to one another, will be reviewed in the following section.

Test-Retest Reliability

Test-retest reliability, as applied to psychometric instruments, refers to the stability of the measuring device over time as well as the temporal stability of the underlying construct (Borg & Gall, 1989; Cronbach, 1960). Test-retest reliability coefficients are correlations between initial and subsequent administrations of the same

measure to the same sample at different time intervals. Anastasi (1988) suggested that test-retest reliability is an integral part of the test development process and “shows the extent to which scores on a test can be generalized over different occasions” (p. 117). Random fluctuations in performance from one testing session to the next can be attributed to several extraneous variables in the testing environment (e.g., weather, lighting conditions, administration procedures) as well as internal and external events in the lives of the test takers such as illness, emotional lability, family disagreements, peer arguments, developmental changes, distress, and fatigue.

While there is general agreement among psychologists that the scores on measures of intelligence, interest, and aptitude should be highly stable over time (i.e., correlation coefficients between the .80s and .90s), it is much less clear whether tests that measure personality or psychopathology should be held to the same standard (Graham, 1993). It is important to differentiate between the *error variance* of scores (fluctuations due to chance factors) and *true variance*, or the actual fluctuations in the construct(s) being measured. However, this is often easier said than done. Flanery (1990) suggested that classic psychometric theory is built on the assumption that the constructs being measured are “trait-like” and relatively stable. Thus, a measure with low test-retest reliability is often judged to be a poor test. However, if a particular construct varies naturally over time, an accurate measurement of such an “unstable” construct will reveal relatively low test-retest reliability coefficients (Anastasi, 1988).

The belief that human behavior is relatively consistent over time is a hotly debated issue (Mischel, 1968). Furthermore, the degree to which emotions and feeling

states change over time is poorly understood. Thus, the determination of an acceptable level of temporal stability on measures of subjective mood states is open to debate. Edelbrock and colleagues (1985) argued that there is no absolute test-retest reliability standard for a child's self-report instrument. However, on objective self-report measures of internalizing constructs such as the RCDS and the CDI, Reynolds (1989) suggested that moderately high test-retest reliability coefficients (.70s and higher) over a period of several weeks are adequate due to the fact that many internalizing symptoms (e.g., depression) fluctuate naturally over time. Kovacs (1992) suggested that "one would not expect a depressive syndrome to remain uniformly stable over months...thus, for a symptom oriented instrument, a two-week test-retest interval may be the most appropriate" (p. 37). Furthermore, when assessing children, the evaluator must remain cognizant of possible developmental changes (e.g., intellectual, social, perceptual, affective) that could lower stability estimates of internalizing disorders (Flanery, 1990). For example, certain early childhood fears and anxieties tend to abate as the child gets older, which has been conceptualized as a normal developmental process (Campbell, 1986; Miller, 1983). In light of these considerations, most researchers recommend brief testing intervals between 2 to 4 weeks (Anastasi, 1988; Borg & Gall, 1989; Kovacs, 1992).

Temporal Stability of Internalizing Disorders

Clarizio (1984) posed the question of whether childhood depression is a chronic, transitory, or recurring condition. According to Clarizio, the answer depends on whether

the sample under investigation is based upon a normal or clinical population, the severity and type of the depression, and how long the sample is followed (i.e., test-retest intervals). Kovacs (1985) argued that children who experience more severe forms of depression often suffer for a significant length of time, up to 5 years or more. However, in a critical review of the epidemiology of childhood depressive disorders, Fleming and Offord (1990) stated that "shortcomings in sampling and considerable inconsistency in the measurement of depression...made it difficult to draw firm conclusions about the prevalence (and persistence) of depression in young people" (p. 571). In addition, determining the temporal stability of internalizing problems such as depression depends upon whether one is measuring depressive symptomology or attempting to determine whether a depressive disorder is present. The specific depressive symptomology can vary over time and the determination of whether a disorder is present may reflect an diverse array of symptom clusters that change in severity above and below a particular diagnostic threshold. Kazdin (1990) described this phenomenon as the distinction between dimensional (symptomology) versus categorical (disorder) assessment.

The temporal stability and persistence of internalizing disorders has been the subject of increased attention in recent years. Early studies characterized internalizing disorders as transient, normal developmental reactions to stress (Achenbach, 1985), and unstable when compared to the relative stability of externalizing disorders (Fischer et al., 1984; Graham & Rutter, 1973; McGee et al., 1985), especially conduct disorder (Offord et al., 1992). However, several recent longitudinal studies have provided evidence to support the idea that internalizing disorders are more persistent than previously thought.

In a study of 436 fourth and fifth graders with depressive symptomology as determined by self-report, peer, and teacher ratings, Tesiny and Lefkowitz (1982) reported that the depressive symptoms were stable over a 5-month interval. McGee and Williams (1988) reported that 31% of depressed 9-year-olds were found to have persistent depressive disorders after 2- and 4-year follow-ups, at the ages of 11 and 13, respectively. Cantwell and Baker (1989) followed a group of children with various internalizing disorders (i.e., depression and anxiety) over a 4- to 5-year period and found that 66% still had internalizing problems at the time of follow-up testing. In another longitudinal study, Nolen-Hoeksema et al. (1992) found that in a large sample of children who initially met criteria for serious levels of depression during initial testing, approximately 40% remained at that level for 6 months to 2 years. Furthermore, in a sample of children between the ages of 4 and 16, Achenbach and his colleagues (Achenbach, Howell, Quay, & Conners, 1991) reported moderate stabilities ($r = .51$) for parent ratings of internalizing symptoms over a 3-year interval (Achenbach et al., 1991).

In another recent longitudinal study, DuBois et al. (1995) investigated the course and stability of self-reported depressive symptoms in a community sample of 435 school-age children. The authors reported that upon initial assessment, 10% of the sample was at or above the recommended cutoff score on the CDI for clinically significant symptoms of depression. Two years later, the authors reported that 32% of the "clinical" sample continued to endorse clinically significant levels of depression and evidenced a greater pattern of impairment across several areas of functioning.

Combined, the results from these studies are promising and will likely contribute

to our understanding of the temporal stability of internalizing disorders in children.

However, there may be differences in the temporal stability of internalizing symptoms when comparing normal and clinical populations of children, and investigators must remain cognizant of this possibility (Finch et al., 1987).

Despite the recent evidence that supports the stability of internalizing disorders, the findings are far from conclusive. It appears that internalizing disorders can be both transient and stable. As previously mentioned, the STAIC has been shown to be effective in differentiating transient (state anxiety scale) anxiety from its more stable counterpart, trait anxiety. In this case, the internalizing problem of anxiety was found to be both unstable and stable across time, depending on the type of anxiety being measured. The differences are reflected in the test-retest reliability coefficients reported for each scale (refer to Table 1). Therefore, attempts to assess the adequacy of test-retest reliability coefficients for measures of internalizing constructs must include a systematic evaluation of the inherent stability of the internalizing constructs in question.

Reliability of Children's Self-Report

As previously mentioned, some researchers have questioned whether young children can reliably report subjective internal states and emotions (Boyle et al., 1993; Costello, 1986; Edelbrock et al., 1985). In a clinical sample of 242 children who were interviewed using the DISC, Edelbrock and colleagues (1985) reported that children below the age of 10 did not give reliable self-reports (average $r = .43$). Boyle and colleagues (1993) reported similar findings in a community sample of children between 6

and 11 years old who were interviewed using the DICA-R (average r for externalizing disorders = .32; average r for internalizing disorders = .06). In both studies, the authors reported consistent age-related increases in the reliability of self-reported symptomology over 1- to 3-week intervals in children 10-12 years of age and older.

While these findings are relevant, it is important to note that both studies incorporated structured interview formats (i.e., DISC, DICA-R) as opposed to objective self-report measures such as the CDI, RCMAS, STAIC, or RCDS. As mentioned above, most structured interview schedules assess whether a particular diagnosis can be either confirmed or disconfirmed; thus the reliability coefficients are based on a particular diagnostic threshold and whether the child reports symptomology consistently (i.e., diagnostic status) across various intervals. In contrast, while many objective self-report measures establish clinical cutoff scores, reliability coefficients are typically calculated by comparing the total scores for the various intervals. As a result, the attenuated reliability coefficients reported for young children using structured interview formats may be attributed to analyses based on discrete variables (diagnostic status) as compared to total score correlations (continuous variables) on objective self-report measures.

In a test-retest reliability study of the CDI, Finch and colleagues (1987) reported that a normal sample of children between the ages of 7 and 12 gave reliable self-reports of depressive symptomology (coefficients ranged from .67 to .82) over short- to medium-length time intervals. The ability of younger children (less than age 10) to reliably report their experience over short- to medium-length intervals on objective self-report measures of internalizing constructs has been established in several studies, including the RCDS

(Reynolds & Graves, 1989), the CDI (Ghareeb & Beshai, 1989; Smucker, Craighead, Craighead, & Green, 1986), the RCMAS (Reynolds, 1981; Wisniewski et al., 1987), and the STAIC (Spielberger, 1973).

Summary

In summary, internalizing disorders in children have been identified as an area of concern for both researchers and clinicians. Internalizing problems have been implicated as a source of significant distress and impairment for a large number of young people. Unfortunately, internalizing problems are typically insidious and often go undetected and untreated. To address this problem of detection, several authors have stressed the importance of using self-report measures in the assessment of internalizing problems, since the children themselves are often the most reliable evaluators of internal states of distress. Currently, there are several psychometrically sound instruments that assess specific internalizing problems and one instrument (YSR) that purports to measure both the internalizing and externalizing broad-band dimensions in children between 11-18 years of age. However, not one of these instruments is designed to assess the broad-band dimension of internalizing symptoms in middle- to late-elementary school-age children even though it has been established that children as young as 8 years old are able to give reliable self-reports of internal states and emotions. In addition, a number of internalizing problems are often comorbid, rendering screening and assessment procedures based upon a single-syndrome instrument incomplete. For the purposes of broad-band screening and assessment as well as the expeditious treatment of childhood internalizing disorders, there

is a need to develop a comprehensive self-report measure of internalizing problems for middle- to late-elementary school-age children. In addition, there is a need to better understand the temporal stability and reliability of self-reported internalizing symptoms in children from a normal population.

CHAPTER III

PURPOSE AND OBJECTIVES

The purpose of this research project was to conduct an investigation of the temporal stability (i.e., test-retest reliability) of self-reported internalizing symptoms in elementary-age children, as measured by the Internalizing Symptoms Scale for Children (ISSC), a self-report instrument currently under development (Merrell & Walters, 1996). The ISSC was designed to measure the broad-band dimension of internalizing problems in middle- to late-elementary school-age children, with the intent of improving clinicians' screening capabilities and their ability to make appropriate decisions regarding the assessment and treatment of internalizing problems.

The objectives of this study were as follows:

1. To determine the test-retest reliability coefficients for the ISSC across 2-, 4-, and 12-week intervals.
2. To determine how the different test-retest intervals affect the magnitude of the stability coefficients.
3. To determine the proportion of the subjects whose ISSC scores are high enough to be considered "at risk" for manifesting internalizing problems (1.5 SDs above the mean).
4. To determine the stability of the subjects' ISSC scores that are high enough during any one of the intervals to be considered "at risk" for manifesting internalizing problems (1.5 SDs above the mean).

Given the aforementioned purpose and objectives, the following four research questions were addressed in this investigation:

1. What is the test-retest reliability of the ISSC across 2-, 4-, and 12-week time intervals? Do the test-reliability coefficients obtained support the instrument's use as a screening and assessment tool for internalizing symptoms in children?
2. How do different retest intervals affect the magnitude of the test-retest reliability coefficients?
3. What proportion of the subjects' ISSC scores is high enough to be considered "at-risk" for manifesting internalizing problems (1.5 SDs above the mean)?
4. What is the temporal stability of self-reported internalizing symptoms as reported by the children whose level of endorsement was in the "at-risk" range during any one or more of the ISSC administrations? In other words, to what degree are the "at-risk" subjects' ISSC scores fluctuating in and out of the "at-risk" range?

CHAPTER IV

METHODOLOGY

Participants

The participants for this study consisted of middle- to late-elementary school-age children between the ages of 8-12. The accessible sample consisted of all third-, fourth-, and fifth-grade students ($N = 199$) from a small elementary school in an urban area in the Intermountain West. Informed consent was obtained from the parents of 144 children (72.36%). The final sample consisted of 131 children (65.82%; 66 boys, 65 girls). The children were between the ages of 8-12, with a mean age of 9.42 years. The students were from grades three to five, with a mean grade of 3.94. The sample was 86% Caucasian, 8% Hispanic, 4% Asian, 1% African American, and <1% Pacific Islander or Native American. Of this sample, 36% of the students were on qualified free or reduced lunch based on low family income status.

Instrumentation

The Internalizing Symptoms Scale for Children (ISSC) is a 48-item self-report research instrument, designed to assess internalizing problems in children. (Refer to Appendix E.) The rationale for developing the ISSC was based upon the apparent dearth of a comprehensive instrument that assesses the broad-band of internalizing problems in children.

Items for the ISSC were developed based upon the *rational-theoretical* approach

described by Lanyon and Goodstein (1982). Inherent within this approach are the *intuitive* and *content* validation methods. Detailed instrument development information is documented by Walters (1995). During the initial stages of item development, specific behavioral domains were identified from which the specific items would be selected. Based upon a review of the factor-analytic studies of childhood psychopathology (Achenbach & Edelbrock, 1983; Quay, 1986), four domains of internalizing symptoms were selected, including: (a) depression, (b) anxiety, (c) somatic complaints, and (d) social withdrawal. Items representative of each domain were selected based upon a comprehensive review of the childhood psychopathology literature (1980-1994), developmental psychology textbooks, current self-report instruments used to measure specific internalizing constructs (e.g., depression and anxiety), and the *DSM-III-R* (APA, 1984). Upon completion of the initial review, 138 nonoverlapping behavioral descriptors were selected for use in the ISSC. The number of items was reduced to 76 after a second review was conducted to eliminate (a) redundant items, (b) developmentally inappropriate items, (c) items that were too difficult or abstract, and (d) items found to be incongruent with the self-report format.

As part of the formal content validation procedure, the remaining 76 items were disseminated to a panel of 25 professionals, including school psychologists, academic professors of psychology, local psychologists in practice, advanced graduate students, educational specialists, and a pediatrician. Each panelist was asked to rate each item on the basis of age appropriateness, representativeness of the internalizing disorders construct, freedom from cultural or gender bias, and clarity of wording. Each item was

rated on a three-point scale (“poor” to “excellent”) and the panelists were encouraged to provide feedback regarding the appropriateness of the items.

After the items that were rated to be inappropriate were eliminated, the item pool was reduced to 59 items. The remaining 59 items were analyzed for readability, and qualitative feedback was solicited from third-, fourth-, fifth-, and sixth-grade students in various schools (Walters, 1995). Ambiguous items were either reworded or deleted, yielding a 54-item research version. After some additional deletions, the current version of the ISSC contains 48 items. The 48 items on the ISSC were analyzed for readability using the *Spache Primary Reading Formula* (grades 1.3 through 3.9) and the *Dall-Chall Readability Formula* (grades four through college) from the *Readability Analysis* computer program by Gamco Industries, Inc. In addition, an expert in reading assessment provided a qualitative analysis of the ISSC items and provided suggestions to improve the age-appropriateness of the ISSC items (Walters, 1995). The average estimated grade level (i.e., readability) for the 48 ISSC items was determined to be 2.0 (range = 1.0-3.8). Written versus oral presentation of ISSC items does not appear to be a factor in influencing self-reported symptoms (Walters & Merrell, 1995).

The endorsement format of the ISSC is based upon a 4-point Likert scale. Respondents can either disavow or endorse symptoms by circling 0 (“never true”), 1 (“rarely true”), 2 (“sometimes true”), and 3 (“often true”). The authors selected the 4-point scale to reduce the problems associated with the “central tendency effect.”

The preliminary standardization sample includes over 2,200 cases from all of the major geographic regions of the United States. The normative sample has been stratified

to represent the population, both geographically and ethnically. Total scores from the normative sample ranged from 5.64 to 146.00, with a mean of 53.95 and a standard deviation of 19.95 (Walters, 1995). Valid cutoff levels (i.e., criterion-related validity) for clinically significant internalizing symptoms have yet to be established for the ISSC.

An internal consistency coefficient of .90 was reported for the entire sample (Walters, 1995). This figure provides strong evidence of the internal stability of the ISSC items. Item total correlations ranged from .17 to .58. Those items that had correlations with the total score of less than .30 were generally considered not to adequately tap the construct of internalizing disorders (Walters, 1995) and were removed from the ISSC.

In a factor analytic study of the ISSC, a two-factor solution was indicated (Merrell & Crowley, 1996). The first factor, Negative Affect/General Distress, contains items that indicate the presence of specific internalizing symptoms or emotional distress. The second factor, Positive Affect, contains items that denote the absence of internalizing symptoms or the presence of positive affect and cognitions incompatible with emotional distress. These findings are consistent with the work of several researchers who suggested that positive and negative expressions of affectivity are independent components that make unique contributions to the etiology and prevention of internalizing problems such as depression and anxiety (Clark, Beck, & Stewart, 1990; Watson, Clark, & Carey, 1988; Wolfe et al., 1987). Several sample ISSC items are listed by factor in Table 2.

Table 2

Sample ISSC Items, Listed by Factor

Factor 1: Negative Affect/General Distress	Factor 2: Positive Affect
I am shy	I feel cheerful
I worry about things	I feel important
I have bad dreams	I have lots of energy
I worry that I will hurt someone	I do things as well as other kids
I have trouble sleeping	I like the way I look
Lots of things scare me	I do well in school
When there is a problem, it is my fault	I feel happy
It is hard for me to breathe	I like myself

In a recent convergent validity study (Merrell, Anderson, & Michael, in press), the ISSC was compared to three instruments that purport to measure construct(s) within the internalizing domain including the CDI, the RCMAS, and the Internalizing Broad-Band score from the YSR. The general descriptions and psychometric properties of these instruments were reviewed previously. Convergent validity coefficients were obtained by computing the Pearson product-moment correlations between the various instruments. The correlation between the ISSC total score and the CDI total score was .75, indicating that the two instruments measure strongly related, but slightly different constructs. The correlation between the total scores of the ISSC and the RCMAS was .78, also an indication of a moderately strong relationship between the two instruments. The correlation between the ISSC total score and the Internalizing Broad-Band score on the

YSR was .86, which was the strongest relationship found in the study. These results provide evidence that the ISSC is a valid broad-band measure of internalizing symptoms (e.g., depression, anxiety) when compared to other well-researched instruments that purport to measure internalizing constructs.

In an analysis of gender differences on the ISSC, Merrell and Dobmeyer (1996) reported that the mean ISSC total scores for all girls in the standardization sample were significantly higher than the mean scores for all boys. The authors stated that these data show that girls tend to endorse a higher degree of internalizing problems than boys.

Procedure

Before data collection began, approval for this investigation was given by the Institution Review Board (IRB) at Utah State University and school district personnel. (Refer to Appendixes A, B, and C.). Prior to the administration of the ISSC, the parent(s) of each of the 199 child participants targeted for participation were sent a letter describing the study, with an attachment to complete and return to the investigator indicating whether they did or did not give their consent for their child to participate in the study. (Refer to Appendix D.) Children who returned their consent forms (regardless of whether consent was given or denied) received a pencil with various graphic designs. Of the 199 consent forms sent out for review, 173 (86.93%) were returned to the investigator. The parent(s) of 144 children (72.36%) gave their informed consent for their children to participate in the study. Individual child subjects were also given the opportunity to decline participation in the study if they desired, even if their parent(s) had given consent

for their child to participate in the investigation. Child subjects who did not participate in the study were not penalized in any way, and they were given an alternative activity (e.g., homework, reading, drawing) to work on during the ISSC administrations.

The ISSC was initially group-administered to 144 children in February of 1996. The confidentiality of the child subjects was protected by assigning each participant an identification number. Only the participants' identification numbers were written on the ISSC protocols to insure proper tracking across the various intervals. When distributing the protocols prior to each administration, the investigator identified the participants by calling out the name listed on a tear-away tab attached to the upper left-hand corner of the protocol. Once the child received his or her protocol, the tear-away tab was removed, leaving only the child's identification number on the protocol. The children were then asked to record the requested information (date, age, sex, grade) on the protocol, with specific instructions to omit their names.

The ISSC was administered to the child subjects in nine home classrooms between the hours of 8:30 a.m. and 1:30 p.m on either a Monday or a Tuesday between February and May of 1996. Classrooms were divided by grade and the number of child subjects in each classroom ranged from 8-25 students. Prior to the administrations, child subjects were asked to disperse themselves around the classroom to encourage confidential and independent completion of the ISSC protocols. The investigator orally presented the directions and the sample item on the ISSC protocol verbatim before the child subjects were asked to complete the ISSC. Child subjects were encouraged to ask for assistance from the investigator if they did not understand the task or a particular item

on the ISSC protocol. Administration time averaged 10-15 minutes per group.

Subsequent to the initial administration, the ISSC was readministered to the same sample of children at intervals of 2 weeks, 4 weeks, and 12 weeks.

During the course of the study, data obtained from 13 children were not included in the analysis due to incomplete ISSC protocols (more than 3 out of the 48 items missing), illness, or an absence during any one of the four ISSC administrations. The final sample consisted of 131 children who were present for all of the administrations. Only those children ($N = 131$) who completed the ISSC during all four administrations were included in the statistical analysis procedures. Missing data (i.e., unanswered items on individual protocols ≤ 3) were dealt with by incorporating item mean substitutions based on the norms of the national standardization sample.

CHAPTER V

RESULTS

The presentation of results is divided into the following four sections: (a) ISSC descriptive statistics, (b) test-retest reliability of the Internalizing Symptoms Scale for Children (ISSC), (c) analysis of variance with repeated measures, and (d) proportion and temporal stability of “at-risk” cases in 131 elementary-age children.

Descriptive Statistics

The mean total ISSC score, median, standard deviation, range, and variance were calculated for each of the four ISSC administrations. The mean ISSC total score for the sample of 131 children was 53.65 ($SD = 19.51$) during the initial administration. The mean ISSC total scores for the sample were 49.56 ($SD = 22.37$), 47.83 ($SD = 21.88$), and 48.07 ($SD = 21.20$) during the subsequent intervals of 2, 4, and 12 weeks, respectively. Combined, the total scores for the sample on the ISSC during all four administrations ranged from 2 to 109. These data are presented in Table 3.

Across all four intervals, the female participants endorsed a higher level of symptomology (average = 2.4 points) on the ISSC as compared to the male participants. However, the differences between the male and female means were not statistically significant, as indicated by independent means t -test results. These data are summarized in Table 4.

Table 3

Descriptive Statistics for the ISSC Scores of 131 Elementary-Age Children for Each Interval

Interval	Mean ISSC Score	Median	<u>SD</u>	Range	Variance
Initial	53.65	53	19.51	6-98	380.79
2-week	49.56	48	22.37	5-109	500.64
4-week	47.83	46.31	21.88	2-98	479.08
12-week	48.07	49	21.20	4-105	449.48

Table 4

Means of the ISSC Scores Over Repeated Administrations Based on Gender with t-Test Results and Significance Levels

Interval	Male <u>M</u> (66)	Female <u>M</u> (65)	<u>t</u> (129)	<u>p</u>
Initial	52.29	55.02	-.80	.43
2-week	48.01	51.14	-.80	.43
4-week	46.84	48.84	-.52	.60
12-week	47.19	48.97	-.48	.63

Test-Retest Reliability of the ISSC

Test-retest reliability coefficients for this investigation were obtained by computing Pearson product-moment (PPM) correlations between the ISSC scores obtained from a sample of 131 children during the initial, 2-week, 4-week, and 12-week administrations of the ISSC. While there were only four administrations of the ISSC,

different combinations of test-retest intervals (e.g., time in between third and fourth intervals = 10 weeks) were used to compute additional interval correlations. The mean reliability coefficient for the 2-week retest intervals was .87 for the ISSC total score. ISSC reliability coefficients for 4- and 12-week retest intervals were .76 and .74, respectively. The test-retest reliability coefficients for the ISSC total scores are presented in Table 5.

Similarly, test-retest reliability coefficients for the two factor scores on the ISSC were computed based upon PPM correlations between the initial and subsequent factor scores on the ISSC. Correlation coefficients for the first factor score (Negative Affect/General Distress) were .85 at 2 weeks, .73 at 4 weeks, and .70 at 12 weeks. Test-retest reliability coefficients for the second factor score (Positive Affect) were .81 at 2 weeks, .79 at 4 weeks, and .72 at 12 weeks. These results are summarized in Tables 6 and 7.

Table 5

Test-Retest Reliability Coefficients for ISSC Total Scores of 131 Elementary-Age Children

Interval	Initial	2-week	4-week	12-week
Initial	--			
2-week	.84	--		
4-week	.76	.90 (2-week)	--	
12-week	.74	.87 (10-week)	.88 (8-week)	--

Table 6

Test-Retest Reliability Coefficients for ISSC Factor 1 Scores (Negative Affect/General Distress) of 131 Elementary-Age Children

Interval	Initial	2-week	4-week	12-week
Initial	--			
2-week	.81	--		
4-week	.73	.89 (2-week)	--	
12-week	.70	.83 (10-week)	.87 (8-week)	--

Table 7

Test-Retest Reliability Coefficients for ISSC Factor 2 Scores (Positive Affect) of 131 Elementary-Age Children

Interval	Initial	2-week	4-week	12-week
Initial	--			
2-week	.79	--		
4-week	.79	.83 (2-week)	--	
12-week	.72	.85 (10-week)	.84 (8-week)	--

Analysis of Variance With Repeated Measures

An analysis of variance (ANOVA) with repeated measures was conducted to determine the stability of ISSC scores across the four intervals (see Glass & Hopkins, 1996 for further review). The results indicated a significant difference among the mean

scores across the four intervals, $F(3, 390) = 12.31, p < .001$. An examination of the means (Table 4) indicates that the main source of the difference among the means is between the initial and subsequent administrations of the ISSC. The results of the ANOVA with repeated measures are summarized in Table 8.

Proportion and Temporal Stability of “At-Risk” Cases

As previously mentioned, empirically supported clinical cutoff scores (criterion-related validity) have yet to be established for the ISSC. However, common cutoff points for self-report measures used to determine whether the level of endorsement is significant enough to indicate “caseness” range from 1.5 to 2.0 standard deviations above the mean for the sample (Merrell, 1994). Therefore, as a preliminary method of evaluating the severity of internalizing symptomology endorsed by the child subjects in the present study, a total score on the ISSC that was equal to or greater than 1.5 standard deviations

Table 8

Analysis of Variance With Repeated Measures

Source of Variance	Sum of Squares	df	Mean Squares	F
Between subjects	205,277.21	130	1579.06	
Within subjects	32,866.29	393		
Time (Interval)	2,843.51	3	947.84	12.31*
Residual	30,022.78	390	76.98	
Total	238,143.50	523		

*significant at $p < .001$.

(SDs) above the mean was used to indicate whether the children were “at-risk” for internalizing symptoms.

A total of 19 out of the 131 child subjects (14.5%) endorsed a level of symptomology on the ISSC that was equal to or greater than 1.5 standard deviations above the mean during one or more of the ISSC administrations. Of the 19 children in the “at-risk” range, 12 of them were in the third grade with an average age of 8.94 years. Of the subjects in the “at-risk” group, 11 of the 19 children were boys. These results are summarized in Table 9.

The average number of intervals during which the 19 “at risk” children reported a level of symptomology that was equal to or greater than 1.5 standard deviations above the mean was 2.33 out of four intervals. The number of children whose level of endorsement was equal to or greater than 1.5 standard deviations was 12 for time 1, 11 for time 2, 12 for time 3, and 7 for time 4 (refer to Table 10). The percentage of the “at-risk” cases that did not drop below 1.5 standard deviations on each of the intervals was 63% (time 1), 58% (time 2), 63% (time 3), and 37% (time 4).

Of the 19 children who were in the “at-risk” group, 5 children were in the “at-risk” range during only one of the testing intervals, whereas 6 of the children were in the “at-risk” group during half of the intervals. Of the 19 “at-risk” children, 7 were 1.5 standard deviations above the mean during three out of the four intervals, while only 1 child remained in the “at-risk” category during all four of the testing intervals.

By using more liberal criteria (1 SD above the mean) to account for the standard error of measurement as well as other forms of error variance, 11 of the 19 (58%)

Table 9

“At-Risk” Cases in the Sample as Determined by Total ISSC Scores

Case	Initial	2-week	4-week	12-week	# of intervals ≥ 1.5 <u>SD</u> ^a	# of intervals ≥ 1 <u>SD</u> ^b
307	78 ^a	94 ^a	90 ^a	60	2	3
318	92 ^a	109 ^a	85 ^a	78 ^b	3	4
320	85 ^a	76 ^b	84 ^a	76 ^b	2	4
322	75 ^b	86 ^a	86 ^a	85 ^a	3	4
332	64	75 ^b	82 ^a	85 ^a	2	3
338	83 ^a	76 ^b	73 ^b	67	1	3
341	83 ^a	84 ^a	83 ^a	71 ^b	3	4
350	83 ^a	81 ^b	71 ^b	76 ^b	1	4
354	86 ^a	70	67	59	1	1
357	66	75 ^b	83 ^a	91 ^a	2	3
358	89 ^a	86 ^a	92 ^a	79 ^b	3	4
364	98 ^a	105 ^a	98 ^a	81 ^a	4	4
421	79 ^b	80 ^b	89 ^a	73 ^b	1	4
443	80 ^a	83 ^a	92 ^a	77 ^b	2	4
449	71	86 ^a	71 ^b	94 ^a	2	3
452	84 ^a	61	47	58	1	1
457	91 ^a	106 ^a	84 ^a	77 ^b	3	4
533	90 ^a	87 ^a	68	80 ^a	3	3
565	83 ^a	95 ^a	77 ^b	105 ^a	3	4

Total # of children = 19; Average # of intervals ≥ 1.5 SDs = 2.33

^a Denotes the total ISSC score was ≥ 1.5 standard deviations above the mean.

^b Denotes the total ISSC score was ≥ 1 standard deviation above the mean.

Table 10

Level of Symptom Endorsement for “At-Risk” Cases for Each Interval

Level of Endorsement	Time 1	Time 2	Time 3	Time 4
≥ 1.5 <u>SDs</u>	12	11	12	7
≥ 1 <u>SD</u>	4	6	4	8
< 1 <u>SD</u>	3	2	3	4
Total	19	19	19	19

“at-risk” cases remained equal to or greater than 1 standard deviation above the mean during all four of the testing intervals. By utilizing the same criteria, 17 of the 19 (89.5%) children reported a level of symptomology on the ISSC that was equal to or greater than one standard deviation on three or more of the four intervals. These results are summarized in Table 9.

CHAPTER VI

DISCUSSION

The four research questions that were posed in this investigation addressed: (a) the test-retest reliability of the ISSC across short to medium length time intervals, (b) the effect that the different retest intervals had on the magnitude of the test-retest reliability coefficients, (c) the proportion of the subjects in the sample whose total ISSC scores were high enough to be considered at risk for manifesting internalizing problems, and (d) the temporal stability of self-reported internalizing symptoms as reported by the children whose level of endorsement was in the “at-risk” range during any one or more of the ISSC administrations. The following discussion will include a brief review and interpretation of the results followed by a discussion of the limitations of this investigation. Finally, suggestions for future research will be presented followed by a review of the clinical implications of the findings.

Test-Retest Reliability of the ISSC

The primary interest of this investigation was the determination of the degree to which middle- to late-elementary school-age children are able to reliably report symptoms of internalizing problems over short- to medium-length time intervals. Overall, the results from this investigation indicate that children between the ages of 8-12 years of age are able to give consistent self-reports of subjective internal states over time as measured by the ISSC. The 2-week test-retest reliability coefficients for the ISSC total

scores were found to be high, ranging from .84 to .90. After 4 weeks, the test-retest reliability coefficient for the ISSC total score was moderate (.76), and only decreased slightly after 12 weeks to .74. The attenuated reliability coefficients reported for the longer retest intervals (4, 12 weeks) are consistent with the findings from previous studies, which provide support for the notion that a number of internalizing mood states fluctuate naturally over time. Nonetheless, the magnitude of the reliability coefficients for the ISSC over short- to medium-length intervals is strong enough to provide empirical support for the instrument as reliable measure of internalizing symptoms in elementary-age children.

As described in Chapter V, the ISSC was administered four times during the course of the study (initial, 2 weeks, 4 weeks, 12 weeks). However, different combinations of the various intervals yielded not only the aforementioned test-retest correlations, but additional coefficients as well, including another 2-week interval (between 2 weeks and 4 weeks) as well as 8-week and 10-week intervals. The test-retest reliability correlations for these additional intervals were consistently high as well, ranging from .90 at 2 weeks to .87 at 10 weeks. While the 10-week test-retest coefficient (.87) did not evidence as much attenuation as the 12-week interval (.74), this finding might be attributed to the decrease in the overall means between the initial administration (53.65) and the 2-week interval (49.56). In other words, the 12-week test-retest coefficient was essentially a comparison of the relationship between the ISSC scores at time 1 and time 4 ($\bar{M} = 53.65$, $SD = 19.51$ and $\bar{M} = 48.07$, $SD = 21.20$) versus a comparison of time 2 and time 4 scores ($\bar{M} = 49.56$, $SD = 22.37$ and $\bar{M} = 48.07$, $SD =$

21.20), which were more similar overall (refer to the correlation matrices in Tables 4-6).

Consistent with the ISSC total score test-retest correlations, the test-retest reliability coefficients for each of the factor scores were found to be moderate to high. Correlations for the ISSC factor 1 (Negative Affect/General Distress) scores ranged from .89 at 2 weeks to .70 after 12 weeks. Similarly, test-retest reliability coefficients for the ISSC factor 2 scores were moderately stable over time, ranging from .83 at 2 weeks to .72 after 12 weeks.

In general, the test-retest reliability coefficients reported for the ISSC are consistent with, and in some cases superior to the findings from studies investigating the reliability of other objective self-report instruments (CDI, RCDS, RCMAS, STAIC, YSR) of internalizing constructs as described in Chapter II. The test-retest reliability of the ISSC was found to be very stable after short intervals and moderately stable after medium-length intervals. These data strongly support the ISSC as a psychometrically sound assessment device.

Reliability of the ISSC Over Repeated Administrations

For the ISSC national normative group, the mean total score was 53.95 for the standardization sample of over 2,200 children (Merrell & Walters, 1996). Similarly, the mean ISSC score for the sample in this study was 53.65 during the initial assessment. However, after each of the three subsequent administrations of the ISSC, the mean scores dropped an average of 5.26 points during time 2 (49.56), time 3 (47.83), and time 4 (48.07).

While an ANOVA with repeated measures revealed a significant difference among the means for each of the four intervals, an examination of the table of means (Table 4) indicates that the source of the difference is likely between the initial and subsequent administrations of the ISSC (refer to Figure 2). This finding is consistent with data reported from other test-retest reliability studies of instruments that purport to measure internalizing constructs (e.g., Reynolds & Graves, 1989; Finch et al., 1987). Possible interpretations of the attenuated mean ISSC scores during subsequent administrations include an overendorsement of internalizing symptomology by distressed children upon initial testing (Reynolds, 1986), an expected variation in reported symptomology due natural fluctuations in mood over time (Kovacs, 1992), and/or a better understanding of the assessment task during subsequent intervals.

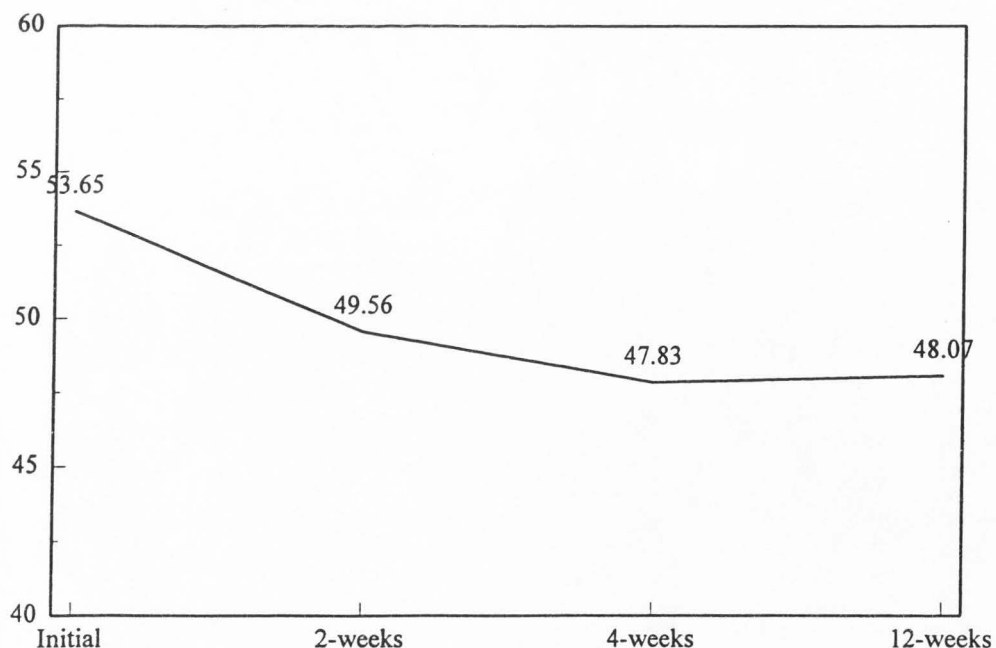


Figure 2. Attenuation of mean ISSC scores across intervals after initial administration.

Temporal Stability of Self-Reported Internalizing Symptoms in the “At-Risk” Group

As was mentioned in Chapter V, of the 131 children in the sample, 19 (14.5 %) endorsed a level of internalizing symptomology on the ISSC that was equal to or greater than 1.5 standard deviations above the mean during one or more of the administrations. Children with this pattern of endorsement were considered to be in the “at-risk” range for internalizing problems. The actual percentage of the sample (14.5%) that endorsed a level of symptomology on the ISSC during any one or more of the intervals that was severe enough to be considered at-risk was within the higher range of base rates (as described in Chapter II) reported for various internalizing problems in normal populations such as depression and anxiety. However, if the ISSC is utilized as an initial screening device as part of a multiple gating assessment procedure whereby a large population is sequentially narrowed down to a smaller population (Merrell, 1994), it appears that a cutoff score of 1.5 standard deviations would minimize the number of false negatives on the ISSC (see Loeber, Dishion, & Patterson, 1984 for further review of this procedure). A higher cutoff score (i.e., 2.0 standard deviations above the mean) on the ISSC might be clinically useful as well, because a smaller percentage of the sample reached the more stringent cutoff point. For example, 6.8% of the sample evidenced a pattern of endorsement that was 2.0 standard deviations above the mean during one or more of the administrations.

On average, the children in the at-risk group were younger than the overall sample

(8.94 years of age) as 12 of the 19 children were in the third grade. While the females in the overall sample tended to consistently endorse a slightly higher level of internalizing symptoms (average = 2.4 points), over half of the children in the at-risk group were males.

Of the children in the at-risk group, 14 (74%) of the 19 children endorsed a level of symptomology that was equal to or greater than 1.5 standard deviations above the mean during at least half of the intervals. Furthermore, if a more liberal at-risk criteria (1 standard deviation above the mean) is used to account for the standard error of measurement as well as other forms of error variance, 17 of the 19 (89.5%) children reported a level of symptomology on the ISSC that was equal to or greater than 1 standard deviation on three or more of the four intervals. These data provide strong support for the temporal stability of self-reported internalizing symptoms as measured by the ISSC.

Limitations and Future Research

Combined with the findings obtained from other studies, the results from this investigation provide a mixture of support and concern regarding the temporal stability of self-reported internalizing symptoms in children between the ages of 8 and 12. While the sample size was reasonably large and representative of this particular region, generalizations made from this study might be limited to populations with similar demographic features since almost 86% of the sample was Caucasian. In addition, the children in this study came from a normal population, even though a subset of the sample

evidenced a level of endorsement on the ISSC that could be considered at-risk for internalizing symptoms. Thus, generalizations regarding the temporal stability of self-reported internalizing symptoms over short- to medium-intervals in a clinical population based on the data obtained in this study are limited.

Future investigations might examine the temporal stability of self-reported internalizing symptoms in clinical and more ethnically diverse populations. In addition, in light of the equivocal nature of the studies that either support or refute the ability of young children (i.e., below the age of 10) to reliably report their internal experience, future investigations might systematically compare those self-report instruments that reportedly have high test-retest reliability (objective self-report) and those that do not (structured interviews).

Clinical Implications of Findings

In summary, the findings from this investigation provide solid empirical support for the idea that children between the ages of 8 and 12 are able to reliably report their experience over short- to medium-length intervals on the Internalizing Symptoms Scale for Children, an objective self-report instrument that purports to measure a broad domain of internalizing problems in elementary-age children. Further examination of the data revealed a possible total score cutoff point of 1.5 standard deviations above the mean for screening purposes used to establish whether the level of endorsement on the ISSC is severe enough to indicate a clinically relevant amount of internalizing symptoms in children.

Overall, the temporal stability of self-reported internalizing symptoms in a sample of 131 elementary school-age children as measured by the ISSC appears to be consistent over time and repeated administrations. In conclusion, these findings provide strong support for the ISSC as a research instrument for screening and assessment of internalizing symptoms in elementary-age children, which may ultimately prove beneficial in the expeditious identification and treatment of childhood internalizing problems.

REFERENCES

- Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor analytic study. Psychological Monographs, 80, (whole No. 615).
- Achenbach, T. M. (1985). Assessment and taxonomy of child and adolescent psychopathology. Beverly Hills, CA: Sage.
- Achenbach, T. M. (1991). Manual for the Child Behavior Checklist/4-18 and 1991 profile. Burlington: University of Vermont Department of Psychiatry.
- Achenbach, T. M., Conners, C. K., Quay, H. C., Verhulst, F., & Howell, C. T. (1989). Replication of empirically derived syndromes as a basis for taxonomy of child and adolescent psychopathology. Journal of Abnormal Child Psychology, 17, 299-323.
- Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: A review and analysis of empirical efforts. Psychological Bulletin, 85, 1275-1301.
- Achenbach, T. M., & Edelbrock, C. S. (1983). Manual for the Child Behavior Checklist and revised behavior profile. Burlington: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., & Edelbrock, C. (1987). Manual for the Youth Self-Report and Profile. Burlington: University of Vermont, Department of Psychiatry.

- Achenbach, T. M., & Edelbrock, C. S. (1989). Diagnostic, taxonomic, and assessment issues. In T. H. Ollendick & M. Hersen (Eds.), Handbook of child psychopathology pp. 53-59). New York: Plenum.
- Achenbach, T. M., Howell, C. T., Quay, H. C., & Conners, C. K. (1991). National survey of competencies and problems among 4- to 16-year-old: Parents' reports for normative and clinical samples. Monographs of Social Research in Child Development, 56, 225.
- Achenbach, T. M., & McConaughy, S. H. (1992). Taxonomy of internalizing disorders of childhood and adolescence. In W. M. Reynolds (Ed.), Internalizing disorders in children and adolescents (pp. 19-60). New York: Wiley.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situation specificity. Psychological Bulletin, 101, 213-232.
- Adams, H. E., & Cassidy, J. F. (1993). The classification of abnormal behavior. In P. B. Sutker & H. E. Adams (Eds.), Comprehensive handbook of psychopathology (pp. 3-25). New York: Plenum.
- American Psychiatric Association. (1980). Diagnostic and statistical manual of mental disorders (3rd ed.). Washington, DC: Author.
- American Psychiatric Association. (1987). Diagnostic and statistical manual of mental disorders (3rd ed., rev.). Washington, DC: Author.
- American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington, DC: Author.

- Anastasi, A. (1988). Psychological testing (6th ed.). New York: Macmillan.
- Anderson, J. C., Williams, S., McGee, R., & Silva, P. A. (1987). DSM-III disorders in preadolescent children. Archives of General Psychiatry, 44, 69-76.
- Angold, A., & Costello, E. J. (1993). Depressive comorbidity in children and adolescents: Empirical, theoretical, and methodological issues. American Journal of Psychiatry, 150, 1779-1791.
- Apolloni, T., & Cooke, T. (1977). Socially withdrawn children: The role of mental health practitioners. Social Behavior and Personality, 5, 337-343.
- Bernstein, G. A., & Borchardt, C. M. (1991). Anxiety disorders of childhood and adolescence. Journal of the American Academy of Child and Adolescent Psychiatry, 30, 519-532.
- Bettes, B., & Walker, E. (1986). Symptoms associated with suicidal behavior in childhood and adolescence. Journal of Abnormal Child and Adolescent Behavior, 14, 591-604.
- Bloom, B. S. (1976). Human characteristics and school learning. New York: McGraw-Hill.
- Borg, W. R., & Gall, M. D. (1989). Educational research: An introduction (5th ed.). New York: Longman.
- Boyle, M. H., Offord, D. R., Racine, Y., Sanford, M., Szatmari, P., Fleming, J. E., & Price-Munn, N. (1993). Evaluation of the Diagnostic Interview for Children and Adolescents for use in general population samples. Journal of Abnormal Child Psychology, 21, 663-681.

- Brady, E. U., & Kendall, P. C. (1992). Comorbidity of anxiety and depression in children and adolescents. Psychological Bulletin, 111, 244-255.
- Breen, T. (1987). Changing family structures. An intervention study in the schools. Unpublished doctoral dissertation, University of Wisconsin-Madison.
- Campbell, S. B. (1986). Developmental issues in childhood anxiety. In R. Gittelman (Ed.), Anxiety disorders of childhood (pp. 24-57). New York: Guilford.
- Cantwell, D. P. (1990). Depression across the early life span. In M. Lewis & S. M. Miller (Eds.), Handbook of developmental psychopathology (pp. 293-309). New York: Plenum.
- Cantwell, D. P., & Baker, L. (1989). Stability and natural history of DSM-III childhood diagnoses. Journal of the American Academy of Child and Adolescent Psychiatry, 28, 691-700.
- Carey, M. P., Finch, A. J., & Imm, P. (1989, November). Differentiating depressive and anxiety disorders from other child psychiatric disorders: The role of overt behavior, cognition, and emotions. Paper presented at the Association for the Advancement of Behavior Therapy, Washington, DC.
- Carlson, G. A., & Garber, J. (1986). Developmental issues in the classification of depressive in children. In M. Rutter, C. E. Izard, & P. B. Read (Eds.), Depression in young people: Developmental and clinical perspectives (pp. 399-434). New York: Guilford.

- Cicchetti, D., & Toth, S. L. (1991). A developmental perspective on internalizing and externalizing disorders. In D. Cicchetti & S. L. Toth (Eds.), Internalizing and externalizing expressions of dysfunction (pp. 1-19). Hillsdale, NJ: Erlbaum.
- Clarizio, H. (1984). Childhood depression: Diagnostic considerations. Psychology in the Schools, 21, 181-197.
- Clark, D., Beck, A., & Stewart, B. (1990). Cognitive specificity and positive-negative affectivity: Complementary or contradictory views of anxiety and depression? Journal of Abnormal Psychology, 99, 148-155.
- Costello, A. J. (1986). Assessment and diagnosis of affective disorders in children. Journal of Child Psychology and Psychiatry, 27, 565-574.
- Costello, E. J. (1989). Developments in child psychiatric epidemiology. Journal of the American Academy of Child and Adolescent Psychiatry, 28, 836-841.
- Costello, E. J., Edelbrock, C., Burns, B. J., Dulcan, M. K., Brent, D., & Janiszewski, S. (1988). Psychiatric disorders in pediatric primary care. Archives of General Psychiatry, 45, 1107-1116.
- Costello, A. J., Edelbrock, L. S., Dulcan, M. K., Kalas, R., & Klaric, S. H. (1984). Report on the NIMH Diagnostic Interview Schedule for Children (DIS-C). Washington, DC: National Institute of Mental Health.
- Cronbach, L. J. (1960). Essentials of psychological testing. (2nd ed.). New York: Harper & Row.
- Cytryn, L., & McKnew, D. H. (1972). Proposed classification of childhood depression. American Journal of Psychiatry, 129, 149-155.

- DuBois, D. L., Felner, R. D., Bartels, C. L., & Silverman, M. M. (1995). Stability of self-reported depressive symptoms in a community sample of children and adolescents. Journal of Clinical Child Psychology, 4, 386-396.
- Edelbrock, C. (1983). Problems and issues in using rating scales to assess child personality and psychopathology. School Psychology Review, 12, 293-299.
- Edelbrock, C. S., Costello, A. J., Dulcan, M. K., Conover, N. C., & Kalas, R. (1986). Parent-child agreement on child psychiatric symptoms assessed via structured interview. Journal of Child Psychology and Psychiatry, 27, 181-190.
- Edelbrock, C. S., Costello, A. J., Dulcan, M. K., & Conover, N. C. (1985). Age differences in the reliability of the psychiatric interview of the child. Child Development, 56, 265-275.
- Finch, A. J., & Rogers, T. R. (1984). Self-report instruments. In T. H. Ollendick & M. Hersen (Eds.), Child behavioral assessment: Principles and procedures. New York: Pergamon Press.
- Finch, A. J., Saylor, C. F., Edwards, G. L., & McIntosh, J. A. (1987). Children's Depression Inventory: Reliability over repeated administrations. Journal of Clinical Child Psychology, 16(4), 339-341.
- Finn, S. E., & Kamphuis, J. H. (1995). What a clinician needs to know about base rates. In J. N. Butcher (Ed.), Clinical personality assessment: Practical approaches (pp. 224-235). New York: Oxford University Press.
- Fischer, M., Rolf, J. E., Hasazi, J. E., & Cummings, L. (1984). Follow-up of a preschool epidemiological sample: Cross-age continuities and predictions of

later adjustment with internalizing and externalizing dimensions of behavior.

Child Development, 55, 137-150.

- Flanery, R. C. (1990). Methodological and psychometric considerations in child reports. In A. M. La Greca (Ed.), Through the eyes of a child: Obtaining self-reports from children and adolescents (pp. 57-82). Boston: Allyn & Bacon.
- Fleming, J. E., & Offord, D. R. (1990). Epidemiology of childhood depressive disorders: A critical review. Journal of the American Academy of Child and Adolescent Psychiatry, 29, 571-580.
- Garfield, S. L. (1993). Methodological problems in clinical diagnosis. In P. B. Sutker & H. E. Adams (Eds.), Comprehensive handbook of psychopathology (pp. 27-46). New York: Plenum.
- Garralda, M. E. (1992). A selective review of child psychiatric syndromes with a somatic presentation. British Journal of Psychiatry, 161, 759-773.
- Ghareeb, G. A., & Beshai, J. A. (1989). Arabic version of the Children's Depression Inventory: Reliability and validity. Journal of Clinical Child Psychology, 18, 323-326.
- Glaser, K. (1967). Masked depression in children and adolescents. American Journal of Psychotherapy, 21, 565-574.
- Glass, G. V., & Hopkins, K. D. (Eds.). (1996). Repeated measures ANOVA. Statistical methods in education and psychology (3rd ed., pp. 572-592). Boston: Allyn & Bacon.

- Graham, J. R. (1993). MMPI-2: Assessing personality and psychopathology (2nd ed.). New York: Oxford University Press.
- Graham, P., & Rutter, M. (1973). Psychiatric disorder in the young adolescent: A follow-up study. Proceedings of the Royal Society of Medicine, 66, 1226-1228.
- Greene, J. W., & Thompson, W. (1984). A physician/psychologist team approach to children and adolescents with recurrent somatic complaints. School Psychology Review, 13, 204-210.
- Harter, S. (1990). Issues in the assessment of the self-concept of children and adolescents. In A. M. La Greca (Ed.), Through the eyes of a child: Obtaining self-reports from children and adolescents (pp. 292-325). Boston: Allyn & Bacon.
- Hodges, K., Kline, J., Fitch, P., McKnew, D., & Cytryn, L. (1981). The Child Assessment Schedule: A diagnostic interview for research and clinical use. Catalog of Selected Documents in Psychology, 11, 56.
- Jensen, P., Roper, M., Fisher, P., Piacentini, J., Canino, G., Richters, J., Rubio-Stipec, M., Dulcan, M., Goodman, S., Davies, M., Rae, D., Shaffer, D., Bird, H., Lahey, B., & Schwab-Stone, M. (1995). Test-retest reliability of the Diagnostic Interview Schedule for Children (DISC 2.1). Archives of General Psychiatry, 52, 61-71.
- Kanshani, J. H., McGee, R. O., Clarkson, S. E., Anderson, J. C., Walton, L. A., Williams, S., Silva, P. A., Robins, A. J., Cytryn, L., & McKnew, D. H. (1983). Depression in a sample of 9 year-old children. Archives of General Psychiatry, 145, 906-954.
- Kauffman, J. M. (1989). Characteristics of behavior disorders of children and youth (4th ed.). Columbus, OH: Merrill.

- Kazdin, A. E. (1988). The diagnosis of childhood disorders: Assessment issues and strategies. Behavioral Assessment, 10, 67-94.
- Kazdin, A. E. (1989). Identifying depression in children: A comparison of alternative selection criteria. Journal of Abnormal Child Psychology, 17, 437-454.
- Kazdin, A. E. (1990). Assessment of childhood depression. In A. M. La Greca (Ed.), Through the eyes of a child: Obtaining self-reports from children and adolescents (pp. 189-233). Boston: Allyn & Bacon.
- Kazdin, A. E., & Petti, T. A. (1982). Self-report and interview measures of childhood and adolescent depression. Journal of Child Psychology and Psychiatry, 23, 437-457.
- Kendall, P. C., Kortlander, E., Chansky, T. E., & Brady, E. U. (1992). Comorbidity of anxiety and depression in youth: Treatment implications. Journal of Consulting and Clinical Psychology, 60, 869-880.
- King, N. J., & Ollendick, T. H. (1989). Children's anxiety and phobic disorders in school settings: Classification, assessment, and intervention issues. Review of Educational Research, 59, 431-470.
- Kovacs, M. (1985). The natural history and course of depressive disorders in children. Psychiatric Annals, 15, 387-389.
- Kovacs, M. (1992). Children's Depression Inventory. New York: Multi-Health Systems.
- Kovacs, M., Feinberg, T., Crouse-Novack, M., Paulauskas, S. L., Pollack, M., & Finkelstein, R. (1984). Depressive disorders in childhood: II. A longitudinal study

- of the risk for subsequent major depression. Archives of General Psychiatry, 41, 643-649.
- Kovacs, M., Goldston, D., & Gatsonis, C. (1993). Suicidal behaviors and childhood-onset depressive disorders: A longitudinal investigation. Journal of the American Academy of Child and Adolescent Psychiatry, 32, 8-20.
- Kurdek, L. A., & Berg, B. (1987). Children's Beliefs About Parental Divorce Scale: Characteristics and concurrent validity. Journal of Consulting and Clinical Psychology, 55, 712-718.
- La Greca, A. M. (1990). Issues and perspectives on the child assessment process. In A. M. La Greca (Ed.), Through the eyes of a child: Obtaining self-reports from children and adolescents (pp. 3-17). Boston: Allyn & Bacon.
- Lanyon, R. I., & Goodstein, L. D. (1982). Personality assessment (2nd ed.). New York: Wiley.
- Last, C. G. (1991). Somatic complaints in anxiety disordered children. Journal of Anxiety Disorders, 5, 125-138.
- Lefkowitz, M. M., & Tesiny, E. P. (1980). Assessment of childhood depression. Journal of Consulting and Clinical Psychology, 48, 43-50.
- Lefkowitz, M. M., & Tesiny, E. P. (1985). Depression in children: Prevalence and correlates. Journal of Consulting and Clinical Psychology, 5, 647-656.
- Loeber, R., Dishion, T. J., & Patterson, G. R. (1984). Multiple gating: A multistage assessment procedure for identifying youths at risk for delinquency. Journal of Research in Crime and Delinquency, 21, 7-32.

- Martin, R. P. (1988). Assessment of personality and behavior problems: Infancy through adolescence. New York: Guilford.
- McGee, R., & Williams, S. (1988). A longitudinal study of depression in nine-year-old children. Journal of the American Academy of Child and Adolescent Psychiatry, 27, 342-348.
- McGee, R., Williams, S., Bradshaw, J., Chapel, J., Robins, A., & Silva, P. (1985). The Rutter Scale for completion by teachers: Factor structure and relationships with cognitive abilities and family adversity for a sample of New Zealand children. Journal of Child Psychology and Psychiatry, 26, 727-739.
- Merrell, K. W. (1994). Assessment of behavioral, emotional, and social problems: Direct and objective methods for use with children and adolescents. White Plains, NY: Longman.
- Merrell, K. W., Anderson, K. E., & Michael, K. D. (1997). Convergent validity of the Internalizing Symptoms Scale for Children with three self-report measures of internalizing problems. Journal of Psychoeducational Assessment, 15, 56-66.
- Merrell, K. W., & Crowley, S. L. (1996). Development and factor structure of a self-report measure for assessing internalizing symptoms in elementary-age children. Manuscript submitted for publication.
- Merrell, K. W., & Dobmeyer, A. C. (1996). An evaluation of gender differences in self-reported internalizing symptoms of elementary-age children. Journal of Psychoeducational Assessment, 14, 196-207.

- Merrell, K. W., Gill, S. J., McFarland, T., & McFarland, H. (1996). Social-emotional adjustment of gifted children on a self-report measure of internalizing symptoms: A comparative study. Psychology in the Schools, 33, 185-191.
- Merrell, K. W., & Walters, A. S. (1996). The Internalizing Disorders Evaluation Scale for Children: Project prospectus and interim users manual. Logan: Department of Psychology, Utah State University.
- Mischel, W. (1968). Personality and assessment. New York: Wiley.
- Miller, L. C. (1983). Fears and anxieties in children. In G. E. Walker & M. C. Roberts (Eds.), Handbook of clinical child psychology (pp. 337-380). New York: Wiley.
- Miller, S. M., Boyer, B. A., & Rodoletz, M. (1990). Anxiety in children. In M. Lewis & S. M. Miller (Eds.), Handbook of developmental psychopathology (pp. 191-207). New York: Plenum.
- Newman, D. L., Moffitt, T. E., Caspi, A., Magdol, L., Silva, P. A., & Stanton, W. R. (1996). Psychiatric disorder in a birth cohort of young adults: Prevalence, comorbidity, clinical significance, and new case incidence from ages 11 to 21. Journal of Consulting and Clinical Psychology, 64, 552-562.
- Nolen-Hoeksema, S., Girgus, J., & Seligman, M. (1992). Predictors and consequences of childhood depressive symptoms: A 5-year longitudinal study. Journal of Abnormal Psychology, 3, 405-422.
- Norvell, N., Brophy, C., & Finch, A. J. (1985). The relationship of anxiety to childhood depression. Journal of Personality Assessment, 49, 150-153.

- Offord, D. R., Boyle, M. H., Racine, Y. A., Fleming, J. E., Cadman, D. T., Blum, H. M., Byrne, C., Links, P. S., Lipman, E. L., MacMillan, H. L., Rae Grant, N. I., Sanford, M. N., Szatmari, P., Thomas, H., & Woodward, C. A. (1992). Outcome, prognosis, and risk in a longitudinal follow-up study. Journal of the American Academy of Child and Adolescent Psychiatry, 31, 916-923.
- Ollendick, T. H., & King, N. J. (1994). Diagnosis, assessment, and treatment of internalizing problems in children: The role of longitudinal data. Journal of Consulting and Clinical Psychology, 62, 918-927.
- Prout, H. T., & Chizik, R. (1988). Readability of child and adolescent self-report measures. Journal of Consulting and Clinical Psychology, 56, 152-154.
- Puig-Antich, J. (1982). The use of RDC criteria for major depressive disorder in children and adolescents. Journal of the American Academy of Child Psychiatry, 21, 291-193.
- Puig-Antich, J., & Chambers, W. J. (1978). The Schedule for Affective Disorders and Schizophrenia for School-Age Children (Kiddie-SADS). New York: New York State Psychiatric Institute.
- Quay, H. C. (1986). Classification. In H. C. Quay & J. S. Werry (Eds.), Psychopathological disorders of childhood (3rd ed., pp. 1-34). New York: Wiley.
- Quay, H. C., & La Greca, A. M. (1986). Disorders of anxiety, withdrawal, and dysphoria. In H. C. Quay & J. S. Werry (Eds.), Psychopathological disorders of childhood (3rd ed., pp. 73-110). New York: Wiley.

- Quay, H. C., & Werry, J. S. (Eds.). (1986). Psychopathological disorders of childhood (3rd ed.). New York: Wiley.
- Rao, U., Weissman, M. M., Martin, J. A., & Hammond, R. W. (1993). Childhood depression and risk of suicide: A preliminary report of a longitudinal study. Journal of the American Academy of Child and Adolescent Psychiatry, 32, 21-27.
- Rehm, L. P., & Tyndall, C. I. (1993). Mood disorders: Unipolar and bipolar. In P. B. Sutker & H. E. Adams (Eds.), Comprehensive handbook of psychopathology (pp. 235-261). New York: Plenum.
- Reich, W., & Welner, Z. (1988). Revised version of the Diagnostic Interview for Children and Adolescents (DICA-R). St. Louis, MO: Department of Psychiatry, Washington University School of Medicine.
- Reynolds, C. R. (1981). Long-term stability of scores on the Revised-Children's Manifest Anxiety Scale. Perceptual and Motor Skills, 53, 702.
- Reynolds, C. R., & Richmond, B. O. (1985). Revised Children's Manifest Anxiety Scale manual. Los Angeles: Western Psychological Services.
- Reynolds, W. M. (1986). A model for screening and identification of depressed children and adolescents in school settings. Professional School Psychology, 1, 117-129.
- Reynolds, W. M. (1989). Reynolds Child Depression Scale. Odessa, FL: Psychological Assessment Resources.
- Reynolds, W. M. (1990). Introduction to the nature and study of internalizing disorders in children and adolescents. School Psychology Review, 2, 137-141.

- Reynolds, W.M. (1992a). The study of internalizing disorders in children and adolescents. In W. M. Reynolds (Ed.), Internalizing disorders in children and adolescents (pp. 1-18). New York: Wiley.
- Reynolds, W. M. (1992b). Depression in children and adolescents. In W.M. Reynolds (Ed.), Internalizing disorders in children and adolescents (pp. 149-253). New York: Wiley.
- Reynolds, W. M. (1992c). Internalizing disorders in children and adolescents: Issues and recommendations for future research. In W.M. Reynolds (Ed.), Internalizing disorders in children and adolescents (pp. 311-317). New York: Wiley.
- Reynolds, W. M., & Graves, A. (1989). Reliability of children's reports of depressive symptomology. Journal of Abnormal Child Psychology.
- Rutter, M. (1989). Isle of Wight revisited: Twenty-five years of child psychiatric epidemiology. Journal of the American Academy of Child and Adolescent Psychiatry, 28, 633-653.
- Sanders, D. E. (1996). The Internalizing Symptoms Scale for Children: A validity study with urban, African-American, seriously emotionally disturbed and regular education students. Unpublished doctoral dissertation, James Madison University, Harrisburg, VA.
- Saylor, C. F., Finch, A. J., Furey, W., Baskin, C. H., Furey, W., & Kelly, M. M. (1984). Construct validity for measures of childhood depression: Application of multitrait-multimethod methodology. Journal of Consulting and Clinical Psychology, 52, 955-967.

- Schaffer, D., Schwab-Stone, M., & Fisher, P. (1993). The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): I. Preparation, field testing, interrater reliability, and acceptability. Journal of the American Academy of Child and Adolescent Psychiatry, 32, 643-650.
- Schwab-Stone, M., Fisher, P., Piacentini, J., Schaffer, D., Davies, M., & Briggs, M. (1993). The Diagnostic Interview Schedule for Children-Revised Version (DISC-R): II. Test-retest reliability. Journal of the American Academy of Child and Adolescent Psychiatry, 32, 651-657.
- Silver, L. B. (1988). The scope of the problem in children and adolescents. In J. G. Looney (Ed.), Chronic mental illness in children and adolescents (pp. 39-51). Washington, DC: American Psychiatric Press.
- Smith, K. (1992). Suicidal behavior in children and adolescents. In W.M. Reynolds (Ed.), Internalizing disorders in children and adolescents (pp. 255-282). New York: Wiley.
- Smith, M. S., Mitchell, J., McCauley, E. A., & Calderon, R. (1990). Screening for anxiety and depression in an adolescent clinic. Pediatrics, 85(3), 262-266.
- Smucker, M. R., Craighead, W. E., Craighead, L. W., & Green, B. J. (1986). Normative and reliability data for the Children's Depression Inventory. Journal of Abnormal Child Psychology, 14, 25-39.
- Spielberger, C. D. (1972). Anxiety as an emotional state. In C. D. Spielberger (Ed.), Anxiety: Current trends in theory and research. New York: Academic.

- Spielberger, C. D. (1973). State-Trait Anxiety Inventory for Children. Palo Alto, CA: Consulting Psychologists Press.
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. D. (1970). STAI: Manual for the State-Trait Anxiety Inventory. Palo Alto: Consulting Psychologists Press.
- Stedman's medical dictionary. (1995). Boston, MA: Houghton Mifflin.
- Stone, W. L., & Lemanek, K. L. (1990). Developmental issues in children's self-reports. In A. M. La Greca (Ed.), Through the eyes of a child: Obtaining self-reports from children and adolescents (pp. 18-56). Boston: Allyn & Bacon.
- Strauss, C. C. (1990). Anxiety disorders of childhood and adolescence. School Psychology Review, 2, 142-157.
- Strauss, C. C., Last, C. G., Hersen, M., & Kazdin, A. E. (1988). Association between anxiety and depression in children and adolescents with anxiety disorders. Journal of Abnormal Child Psychology, 16, 57-68.
- Taylor, J. A. (1951). The relationship of anxiety to the conditioned eyelid response. Journal of Experimental Psychology, 41(2), 81-92.
- Tesiny, E. P., & Lefkowitz, M. M. (1982). Childhood depression: A 6-month follow-up study. Journal of Consulting and Clinical Psychology, 50, 778-780.
- Walker, L. S., & Greene, J. W. (1989). Children with recurrent abdominal pain and their parents: More somatic complaints, anxiety, and depression than other patient families? Journal of Pediatric Psychology, 14, 231-243.

- Walters, A. S. (1995). Development and preliminary validation of a self-report instrument for assessing internalizing problems in children. Unpublished doctoral dissertation, Utah State University, Logan.
- Walters, A. S., & Merrell, K. M. (1995). Written versus oral administration of a social-emotional self-report test for children: Does method of execution make a difference? Psychology in the Schools, *32*, 186-189.
- Watson, D., Clark, L. A., & Carey, G. (1988). Positive and negative affectivity and their relation to anxiety and depressive disorders. Journal of Abnormal Psychology, *97*, 346-353.
- Weiss, B., Wiesz, J. R., Politano, M., Carey, M., Nelson, W. M., & Finch, A. J. (1991). Developmental differences in the factor structure of the Children's Depression Inventory. Psychological Assessment, *3*(1), 38-45.
- Weissman, M. M., Wickramaratne, P., Warner, V., John, K., Prusoff, B. A., Merikangas, K. R., & Gammon, D. (1987). Assessing psychiatric disorders in children: Discrepancies between mothers' and children's reports. Archives of General Psychiatry, *44*, 747-753.
- Werry, J. S. (1986). Physical illness, symptoms, and allied disorders. In H. C. Quay & J. S. Werry (Eds.), Psychopathological disorders of childhood (3rd ed., pp. 232-293). New York: Wiley.
- Wisniewski, J. J., Mulick, J. A., Genshaft, J. L., & Coury, D. L. (1987). Test-retest reliability of the Revised Children's Manifest Anxiety Scale. Perceptual and Motor Skills, *65*, 67-70.

Wolfe, V. V., Finch, A. J., Saylor, C. F., Blount, R. L., Pallmeyer, T. P., & Carek, D. J.

(1987). Negative affectivity in children: A multitrait-multimethod investigation.

Journal of Consulting and Clinical Psychology, 55, 245-250.

APPENDIXES

Appendix A:

Statement of the Principal Investigator to the
Institutional Review Board for Proposed
Research Involving Human Subjects

**Statement of the PI to the IRB for Proposed
Research Involving Human Subjects**

Proposal Title An Investigation of the Temporal Stability of Self-Reported
Internalizing Symptoms in Elementary-Age Children

Principal Investigator* Kenneth W. Merrell Dept. Psych UMC 494 Ext. 72034

Student Researcher Kurt D. Michael Dept. Psych UMC 473 Ext. 73059

- A. Human subjects will participate in this research and be asked to do the following: Complete a 54-item self-report test of internalizing symptoms that will take approximately 20 minutes to administer (each administration).
- B. The potential benefits to be gained from the proposed research are: 1) New understanding of the temporal stability of internalizing symptoms over time; 2) establishment of test-retest reliability of the Internalizing Disorders Evaluation Scale for Children (IDESC).
- C. The risk(s) to the rights and welfare of human subjects involved are: No risks are apparent.
- D. The following safeguards/measures to mitigate/minimize the identified risks will be taken: Parents will have the opportunity to decline their child's participation in the study and the children themselves will have the opportunity to decline participation as well.
- E. The informed consent procedures for subjects will be as follows: (Explain procedures to be followed and attach an example of the informed consent instrument) A letter will be sent to the parents of each potential subject explaining the study and giving them the opportunity to decline participation.
- F. The following measures regarding confidentiality of subjects will be taken: The data will be coded so that the subjects will not be personally identified.
- G. Other: (If, in your opinion no, or minimal, risk to subjects exists, please explain in this section) There is no empirical evidence that completing a social-emotional measure poses any risks or danger to child subjects.

Kenneth W. Merrell
Principal Investigator Signature*

Kurt D. Michael
Student Researcher Signature

* A student researcher should name his/her advisor or chairman as the principal investigator. Both are required to sign this form.

Appendix B:

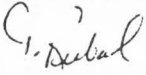
Institutional Review Board Letters of Approval



VICE PRESIDENT FOR RESEARCH OFFICE
Logan, Utah 84322-1450
Telephone: (801) 797-1180
FAX: (801) 797-1367
INTERNET: [pgerity@champ.usu.edu]

February 14, 1996

TO: Dr. Ken Merrell-PI
Kurt Michael-Student Researcher

FROM: True Rubal 

SUBJECT: "An Investigation of the Temporal Stability of Self-Reported
Internalizing Symptoms in Elementary-Age Children"

This protocol was reviewed and approved by the IRB on 2 February 1996 pending a revised Informed Consent. Our office received the revised consent on 14 February 1996. You may consider this to be your approval for your study.

If there should be any changes in this protocol as to methodology etc., it will need to be resubmitted to the IRB. A status report (continuing review) will be due one year from the approval date. Also, please keep the committee advised of any changes, adverse reactions or termination of this study. Thank you.


Utah State
UNIVERSITY

VICE PRESIDENT FOR RESEARCH OFFICE
Logan, Utah 84322-1450
Telephone: (801) 797-1180
FAX: (801) 797-1367
INTERNET: [pgerity@champ.usu.edu]

February 19, 1996

MEMORANDUM

TO: Ken Merrell
Kurt Michael

FROM: True Rubal, Secretary to the IRB 

SUBJECT: An Investigation of the Temporal Stability of Self-Reported Internalizing
Symptoms in Elementary-Age Children

The above referenced proposal was reviewed and approved by the IRB on February 14, 1996.
You may consider this letter to be your approval for your study.

Any deviation from this protocol will need to be resubmitted to the IRB. This includes any
changes in the methodology of procedures in this protocol. A study status report (stating the
continuation or conclusion of this proposal) will be due in one year from the date of this letter.

Please keep the committee advised of any changes, adverse reactions or the termination of this
study. I can be reached at x71180.

Appendix C:

Permission Letter from the

Ogden City School District

OGDEN CITY SCHOOL DISTRICT
Department of Special Education and Student Services

MEMO

TO: Beverly Wilcox, Director, Special Education
Dale Thompson, Principal, Hillcrest

FROM: Cher King, Psychologist *ck*

DATE: January 17, 1995

RE: School psychology practicum student research project

As you know, Kurt Michael from Utah State University is doing his school psychology practicum with us this school year. He is currently working on his master's thesis, which involves helping to gather reliability data for a new internalizing disorders self-report inventory. His work at the university is being supervised by Dr. Ken Merrell, and I am his field supervisor for his work here in Ogden at the present time.

Kurt's project has been approved by his thesis committee and is presently being reviewed by the university's Institutional Review Board (which must approve all research projects involving human subjects). He anticipates the IRB will approve this project by January 26; the IRB has approved an identical project in the recent past; and Kurt would not undertake the project in Ogden unless it is approved.

Enclosed is a copy of the research proposal. The parental consent form (last page) has been changed. It is my recommendation, after conferring with Carol Lehr at the State Office of Education, that an active, rather than passive, consent form be used. In other words, no child would participate in the study without written permission from the parent. The consent form I am proposing is attached to this memo.

Kurt plans to hand out the parental consent form to each third, fourth, and fifth grade student at Hillcrest on Tuesday, January 22. He would like to go into each class and give the consent forms to the students himself with a brief explanation. He would like to ask that the consent forms be returned to the school secretary. Finally, he would like to leave some treats with the secretary to give the students as they turn in their consent forms. The following Tuesday, he will give a second copy to students who have not yet turned in their forms.

Kurt will be administering the self-report inventory to one whole class at a time, i.e., a group administration. This will take approximately 15-20 minutes. He will re-administer the inventory to the same students an additional two or three times, depending upon which group they are in, before the end of school.

Please contact me if you have any questions or concerns about any aspect of this project. If you would like to speak directly with Kurt, his number at Utah State is (801) 797-3059.

Appendix D:

Parent Letter and Informed Consent Form

OGDEN CITY SCHOOL DISTRICT
 Department of Special Education and Student Services
 1950 Monroe Blvd., Ogden, UT 84401

January 23, 1996

Dear Hillcrest Parent:

We are seeking permission from the parents of all the students in grades three, four, and five for these students to participate in a screening activity for the purpose of developing a new psychological test for children. This new test will ultimately be used to help identify symptoms of depression and anxiety in children in grades three through six. The purpose of this particular study will be to see whether children's responses remain consistent over time, i.e., to see if they tend to answer the same questions the same way across several different screenings.

For the screening activity, the students in each classroom will be asked to take about 15-20 minutes to respond to a number of questions regarding their mood, the way they feel about themselves, and certain behaviors they may display that are related to depression or anxiety. Examples of actual statements in the screening include "I am shy," "I worry about things," "I am cheerful," "I feel very tired," and "I am happy." The children will respond to these items by indicating whether each item is never, hardly ever, sometimes, or often true for them. They will mark their responses to the statements on an answer sheet.

Participation will be completely voluntary. Any student who does not wish to participate will be excused from the activity without consequence, and no child will participate without prior written consent from the parent. We believe there is very minimal risk in this activity, and children who participate will not be personally identified in any way. They will be providing information regarding their age, grade, and gender, but student names will not remain on the answer sheets. Parents may examine the screening instrument if they wish, though copies of the instrument may not be made.

The study is being conducted by Kurt Michael, a graduate student in psychology from Utah State University. Mr. Michael is presently doing a school psychology practicum in this school district. If you have any questions about this activity, please feel free to contact Mr. Michael's district field supervisor, Dr. Cher King, at 625-8729, or the Hillcrest principal, Mr. Dale Thompson, at 625-8805.

We would appreciate your returning this consent form as soon as possible. Please sign and return this form through your student, or in person, or by mail, to Hillcrest. Please return the form whether you do or do not wish your student to participate, so we will be certain of your wishes. Thank you for your support of this very important research project.

Sincerely,

Dale Thompson

Dale Thompson, Principal

Cher King

Cher King, Ph.D., Psychologist

I DO give my permission for my child, _____, to participate in the activity described above. (name of student)

 Signature of parent or guardian

 Date

I DO NOT give my permission for my child, _____, to participate in the study described above. (name of student)

 Signature of parent or guardian

 Date

Appendix E:

ISSC Protocol

ISSC

Date _____ Age _____ Grade _____

I am a: Boy Girl (circle one)

Directions

These sentences tell some ways that boys and girls sometimes feel. Read each sentence and decide how often it is true for you. Ask yourself, "Is this *Never true*, *Hardly Ever true*, *Sometimes true*, or *Often true* for me?" After you have decided how often each sentence is true for you, make an X in the circle that goes with that answer. There are no right or wrong answers, just choose the answer that tells how you feel.

Example

	Never True	Hardly Ever True	Sometimes True	Often True
I feel like reading a book	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How true is this for me?	Never True	Hardly Ever True	Sometimes True	Often True
1. I am shy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. I worry about things	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. I feel cheerful	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. I have bad dreams	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. I feel important	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Things are hard for me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
7. I feel lonely	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
8. I worry that I will hurt someone	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
9. I have lots of energy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
10. I have trouble sleeping	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. I feel dizzy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. I feel upset	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. I believe I am good at lots of things	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14. I feel like I have made too many mistakes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Lots of things scare me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Other kids like me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. I feel like crying	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. When there is a problem it is my fault	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. It is hard for me to breathe	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. I do things as well as other kids	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

How true is this for me?	Never True	Hardly Ever True	Sometimes True	Often True
21. I worry that something bad will happen . . .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
22. I like the way I look	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
23. I feel sad	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. I get scared for no reason	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. My stomach hurts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26. My head hurts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27. I feel sorry for myself	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28. It is hard for me to sit still	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
29. I feel like being alone	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30. It is hard for me to think	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31. I laugh and smile as much as other kids . . .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32. My feelings get hurt easily	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33. Nothing is fun for me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. I have a hard time making up my mind . . .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35. I think about hurting myself	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36. I do well in school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37. It seems like no one cares about me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
38. I feel happy	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. I feel very tired	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. I don't feel like doing anything	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41. I like myself	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42. I worry that other people will not like the way I do things	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43. I hate it when I am the center of attention .	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. Bad things happen to me	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45. I think about dying	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46. My hands and feet feel sweaty	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47. I feel like playing with my friends	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48. I can't do anything right	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>